

Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/107800/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Forootan, Ehsan ORCID: <https://orcid.org/0000-0003-3055-041X>, Kusche, Jürgen, Talpe, Matthieu, Shum, C. K. and Schmidt, Michael 2018. Developing a Complex Independent Component Analysis (CICA) technique to extract non-stationary patterns from geophysical time series. *Surveys in Geophysics* 39 (3), pp. 435-465. 10.1007/s10712-017-9451-1 file

Publishers page: <http://dx.doi.org/10.1007/s10712-017-9451-1>
<<http://dx.doi.org/10.1007/s10712-017-9451-1>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Developing a Complex Independent Component Analysis (CICA) Technique to Extract Non-stationary Patterns from Geophysical Time Series

Ehsan Forootan¹  · Jürgen Kusche² · Matthieu Talpe³ · C. K. Shum^{4,5} · Michael Schmidt⁶

Received: 25 May 2017 / Accepted: 16 November 2017

© The Author(s) 2017. This article is published with open access at Springerlink.com

Abstract In recent decades, decomposition techniques have enabled increasingly more applications for dimension reduction, as well as extraction of additional information from geophysical time series. Traditionally, the principal component analysis (PCA)/empirical orthogonal function (EOF) method and more recently the independent component analysis (ICA) have been applied to extract, statistical orthogonal (uncorrelated), and independent modes that represent the maximum variance of time series, respectively. PCA and ICA can be classified as stationary signal decomposition techniques since they are based on decomposing the autocovariance matrix and diagonalizing higher (than two) order statistical tensors from centered time series, respectively. However, the stationarity assumption in these techniques is not justified for many geophysical and climate variables even after removing cyclic components, e.g., the commonly removed dominant seasonal cycles. In this paper, we present a novel decomposition method, the complex independent component analysis (CICA), which can be applied to extract non-stationary (changing in space and time) patterns from geophysical time series. Here, CICA is derived as an extension of real-valued ICA, where (a) we first define a new complex dataset that contains the observed time series in its real part, and their Hilbert transformed series as its imaginary part, (b) an

✉ Ehsan Forootan
ForootanE@cardiff.ac.uk

¹ School of Earth and Ocean Sciences, Cardiff University, Main Building, Park Pl, Cardiff CF10 3AT, UK

² Institute of Geodesy and Geoinformation, University of Bonn, Nußallee 17, 53115 Bonn, Germany

³ Aerospace Engineering Sciences, University of Colorado Boulder, 2598 Colorado Ave, Boulder, CO 80302, USA

⁴ Division of Geodetic Science, School of Earth Sciences, Ohio State University, 275 Mendenhall Lab, 125 South Oval Mall, Columbus, OH 43210, USA

⁵ State Key Laboratory of Geodesy and Earth's Dynamics, Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Wuhan 430077, China

⁶ German Geodetic Research Institute, Technical University of Munich, Arcisstraße 21, 80333 Munich, Germany

ICA algorithm based on diagonalization of fourth-order cumulants is then applied to decompose the new complex dataset in (a), and finally, (c) the dominant independent complex modes are extracted and used to represent the dominant space and time amplitudes and associated phase propagation patterns. The performance of CICA is examined by analyzing synthetic data constructed from multiple physically meaningful modes in a simulation framework, with known truth. Next, global terrestrial water storage (TWS) data from the Gravity Recovery And Climate Experiment (GRACE) gravimetry mission (2003–2016), and satellite radiometric sea surface temperature (SST) data (1982–2016) over the Atlantic and Pacific Oceans are used with the aim of demonstrating signal separations of the North Atlantic Oscillation (NAO) from the Atlantic Multi-decadal Oscillation (AMO), and the El Niño Southern Oscillation (ENSO) from the Pacific Decadal Oscillation (PDO). CICA results indicate that ENSO-related patterns can be extracted from the Gravity Recovery And Climate Experiment Terrestrial Water Storage (GRACE TWS) with an accuracy of 0.5–1 cm in terms of equivalent water height (EWH). The magnitude of errors in extracting NAO or AMO from SST data using the complex EOF (CEOF) approach reaches up to $\sim 50\%$ of the signal itself, while it is reduced to $\sim 16\%$ when applying CICA. Larger errors with magnitudes of $\sim 100\%$ and $\sim 30\%$ of the signal itself are found while separating ENSO from PDO using CEOF and CICA, respectively. We thus conclude that the CICA is more effective than CEOF in separating non-stationary patterns.

Keywords Independent component analysis (ICA) · Complex ICA (CICA) · Time series analysis · Signal separation · Non-stationary decomposition · Terrestrial water storage (TWS) · Sea surface temperature (SST)

1 Introduction

Geophysical and climatological observations, such as the time series of global terrestrial water storage (TWS, Tapley et al. 2004), sea level (Shum and Kuo 2010), and sea surface temperature (SST, Reynolds et al. 2002), contain many inherent time scales, which reflect the complex processes that cause their variations. Traditionally, parametric methods such as regression techniques have been applied to analyze these observations, for which one assumes that the observed time series consists of different parts, for example, a trend (defined as long-term evolution of the series), periodic components including seasonal cycles, and a random part, i.e., noise. Multivariate linear regression (MLR) is a common technique to perform such analysis (Rencher and Christensen 2012). Each part of the model is then accounted for by introducing pre-defined base functions. Finally, the sought-for parameters that are coefficients of the base functions are approximated using, e.g., a least squares adjustment (Koch 1999).

Selecting appropriate base functions to meaningfully represent the behavior of observations is a difficult task in parametric techniques. For example, several studies indicate that the long-term variability of climate records is not perfectly linear in time. Their periodical components cannot be adequately explained by sinusoids, for example, see time series of TWS in Schmidt et al. (2008) and also see discussions on the modulation of amplitudes in SST time series (Moore et al. 2017). Furthermore, it is very difficult to detect whether a trend in these time series is significant or whether it is part of an oscillation (Matalas 1997). Therefore, statistical methods that extract data-adjusted spatial and

temporal patterns from observations have garnered increased interest (von Storch and Zwiers 1999).

Several principles have been developed in statistics to extract linear and/or nonlinear parameterizations of random variables. The term ‘statistical decomposition’ is applied for ‘transforming’ or ‘separating’ multivariate sampled variables (e.g., observed geophysical time series or model simulations) into ‘mathematical components’, which is also known as ‘statistical modes’ (Preisendorfer 1988). The algorithms that are used in the statistical techniques to find such parameterizations can be categorized according to the statistical information used in their decomposition procedure, for example, (a) ‘second-order’ and (b) ‘higher-order’ techniques (Cardoso 1999; Hyvärinen 1999a). They can also be classified, based on how the statistics are estimated, into (A) ‘stationary’ and (B) ‘non-stationary’ techniques. Decomposition techniques have also been discussed under the ‘blind source separation (BSS)’ theme, which aims at recovering unobserved patterns or ‘sources’ from observations that are a ‘mixture’ of these sources (in the presence of noise) measured by an array of sensors (Hyvärinen and Oja 2000). In other words, the term ‘data matrix (observations)’ used in decomposition techniques is equivalent with the ‘mixture (matrix)’ in BSS, and ‘statistical modes’ are equivalent with the terms ‘source(s)’ and ‘(de)mixing matrix’ used in the BSS techniques. The BSS view has been applied in many disciplines including computer science and feature recognition (e.g., Liu and Wechsler 2003), biomedical sciences (e.g., James and Hesse 2005), brain imaging (e.g., Anemüller et al. 2003, 2004; Jung et al. 2005), and many other examples.

In general, second-order decomposition methods (a) try to find the statistical modes (abbreviated as ‘modes’ henceforth) using only the information contained in the autocovariance or autocorrelation matrices, built on the observed time series. Therefore, the first-order statistical moments, i.e., mean values, and then second-order moments, i.e., covariances, are used in (a). Higher-order decomposition methods (b) go one step further than (a) by incorporating higher than two statistical moments (e.g., measures of statistical skewness and kurtosis) in their procedure (Hyvärinen 1999b). Therefore, methods in (a) assume that the statistical moments of up to the second order adequately represent the probability distribution of observations, while those of (b) are applied when the probability distribution of time series is non-Gaussian. In this case, more statistical moments are needed to represent the underlying distribution of the observations (see details in Forootan 2014, chapters 3 and 4).

By definition, a stationary process (A) corresponds to a situation in which the joint probability distribution of variables (time series) does not change with time (Priestley 1988). In contrast, for non-stationary processes (B), the statistical measures (e.g., mean, variance and higher-order statistical moments) change with time. The physical interpretation of (B) is that the observations are associated with phenomena with a shape (extension) and/or strength that evolves in time. This is the case for many geophysical time series; for example, by looking at the global hydrological water fluxes, one can see the amplitude of seasonal cycles as well as their spread change in time (Eicker et al. 2016). This can also be detected in longer time series such as precipitation, sea surface temperature, and sea surface pressure (Hannachi et al. 2007; Timm et al. 2005), which reflect the dynamic of spatially and temporally variable phenomena such as those related to the El Niño Southern Oscillation (ENSO, Trenberth 1990), the North Atlantic Oscillation (NAO, Feldstein 2003), and the Indian Ocean Dipole (IOD, Saji et al. 1999; Krishnaswamy et al. 2015). Generally speaking, based on how the statistical information is computed in (a) and (b), the techniques can potentially deal with stationary (A) and non-stationary (B) property of time series.

In the following paragraphs, common eigenspace techniques that are widely used in climate, geophysics, and hydrology research for signal decomposition are introduced and classified in the (a), (b), (A), and (B) categories. Our motivation to introduce a new decomposition method is also justified.

Principal component analysis (PCA), also called empirical orthogonal function (EOF, Preisendorfer 1988), is among the most popular second-order analysis techniques, therefore classified as (a), and often used to extract dominant orthogonal modes from datasets in various disciplines (see, e.g., Wallace et al. 1992; Fenoglio-Marc 2001; Wouters and Schrama 2007; Omondi et al. 2013). More recently, the higher-order statistical technique of independent component analysis (ICA, Cardoso and Souloumiac 1993; Hyvärinen 1999a, classified here as (b)) has been introduced in order to decompose these data into statistically independent components (e.g., Aires et al. 2002; Westra et al. 2007; Hannachi et al. 2009; Frappart et al. 2010, 2011). Forootan and Kusche (2012, 2013) argue that different physical processes generate statistically independent source signals that are superimposed in geophysical time series; thus, application of ICA likely helps separating (extracting) their contribution from the total signal. Therefore, in the recent studies (e.g., Forootan et al. 2012; Awange et al. 2014; Boergens et al. 2014; Gualandi et al. 2016; Ming et al. 2016), ICA has been preferred over the ordinary extensions of the PCA/EOF approach, such as the rotated EOF (REOF) technique applied in, e.g., Richman (1986) and Lian and Chen (2012).

PCA and ICA (respectively a and b as defined above) are stationary techniques. This means that for PCA, the autocovariance matrix or autocorrelation matrix (see, e.g., Preisendorfer 1988) is used to estimate the orthogonal (statistically uncorrelated) modes. For ICA, the diagonalizing higher (than second)-order statistical tensor (Cardoso and Souloumiac 1993; Forootan and Kusche 2012) or a measure of non-Gaussianity (Hyvärinen 1999b; Boergens et al. 2014) is used to estimate the independent modes. The mentioned ICA criteria are formulated with the fundamental assumption that the estimated statistics (cumulants or non-Gaussianity measures) do not evolve in time, i.e., the stationary assumption. Although both PCA and ICA techniques are efficient in separating signals with various temporal behaviors, they cluster out-of-phase variability of time series as demonstrated in Horel (1984) and Forootan (2014).

As a result, the ordinary PCA approach has been modified to better deal with non-stationary information, which yielded methods such as the extended empirical orthogonal function (EEOF, Weare and Nasstrom 1982) and the complex empirical orthogonal function (CEOF, Rasmusson et al. 1981). EEOF is also called multi-channel singular spectrum analysis (MSSA, Broomhead and King 1986a, b). Non-stationary is introduced in these techniques by incorporating time and/or space lag information while estimating statistical moments (for more details, see, e.g., Hannachi et al. 2007). Various applications indicate a better performance of these extensions when extracting non-stationary behaviors in a few dominant modes (see, e.g., Rangelova et al. 2012; Forootan et al. 2016).

In this study, the ICA technique is extended to deal with non-stationarity of geophysical time series, similarly to how CEOF extends PCA. This has been done by generating a new dataset that contains the observed time series in its real part. The out-of-phase patterns of these time series are estimated by applying a Hilbert transformation (Horel 1984) and are considered to be the imaginary part of the new dataset. The Hilbert transformation shifts the observed time series by 90° in the frequency domain and therefore introduces information about (an approximation of) the rate of change of original time series in the decomposition process (see Appendix 1). The derived complex dataset is used in the ICA procedure of Forootan and Kusche (2012) to extract the dominant independent space and

time amplitudes and associated phase propagations. This new extension of the ICA method is called ‘complex ICA (CICA)’ in this paper.

It is worth mentioning that different criteria exist, which can be used to measure mutual independence of sources and equivalently for implementing ICA/CICA. For example, Fu et al. (2015) argue that three properties, i.e., non-Gaussianity, non-whiteness, and non-circularity, are implemented in most of the ICA algorithms to approximate statistical independence. Given the fact that considering only one of these properties might not be sufficient to separate sources with variety of probability distributions, they introduce a new CICA algorithm that combines these three criteria and illustrate its benefits particularly when sources have proportional covariance matrices. In this study, the main aim is to extract trends, cyclic, and semi-cyclic sources with distinguished frequencies, which avoid the mentioned problem. Therefore, we use the joint approximate diagonalization of eigenmatrices (JADE, Cardoso and Souloumiac 1993, 1995), which is a tensorial approach and is straight forward to be used for estimating the independence of statistical modes (see Sect. 2). The efficiency of JADE in separating cyclic signals is proved in Forootan and Kusche (2013).

An advantage of CICA over the already existing EEOF/MSSA and CEOF techniques is that it incorporates higher-order statistical information, which likely reduces clustering of different physical modes within single extracted ‘mathematical’ modes (see the results in Sects. 5 and 6). It is worth mentioning that most previous CICA algorithms have been defined for random variables that are naturally complex (e.g., Cardoso and Souloumiac 1993). Thus, a distinguished difference of the presented algorithm with existing ones is the transformation from real-valued time series to the complex variables, applying the ICA technique, and finally recovering the independent modes that reconstruct (an approximation of) the introduced real-valued time series. A classification of the methods mentioned above into a, b, A, and B is provided in Table 1.

After briefly reviewing the mathematical derivation of CICA, we focus on assessing the skill of CICA for extracting relevant information from for two geophysically meaningful applications. First, we use time series of TWS from the Gravity Recovery And Climate Experiment (GRACE, Tapley et al. 2004) mission (March 2002 onwards). GRACE TWS data represent integrated changes in all forms of water storage above and underneath the surface of the Earth, i.e., the sum of groundwater, soil moisture and permafrost, surface water, snow/ice, and biomass. These changes cause anomalies of different time scales that must be separated to allow their interpretation. Here, we assess whether it is possible to isolate the long-term linear trends in TWS from seasonal changes alongside semi-cyclic

Table 1 Classification of the decomposition techniques with respect to the statistical information used in their process

	Second order (a)	Higher order (b)
Stationary (A)	PCA/EOF (e.g., Preisendorfer 1988)	ICA (e.g., Forootan and Kusche 2012, 2013) REOF (e.g., Richman 1986)
Non-stationary (B)	EEOF/MSSA (e.g., Broomhead and King 1986a, b) CEO (e.g., Rasmusson et al. 1981)	CICA (this study)

variability due to the dominant influence of the ENSO. Therefore, through a careful synthetic study, we will justify the application of CICA for extracting ENSO and non-ENSO modes from GRACE TWS data similar to Eicker et al. (2016).

The second application of CICA involves the analysis of SST data over the Atlantic and Pacific Oceans. Using this example, we will show whether complicated non-stationary variations such as those related to NAO can be separated from the Atlantic Multi-decadal Oscillation (AMO). The separation of ENSO and the Pacific Decadal Oscillation (PDO) is also discussed.

The contribution of this study is threefold: (1) CICA is mathematically defined, and the details of estimating associated complex space and temporal components are described. (2) The skill of CICA is compared to CEOF and ordinary PCA in a realistic simulated environment, where the true mathematical modes are known by definition. This simulation specifically addresses the separation of a trend from seasonal and low-frequency climate-driven patterns, here due to ENSO, in the presence of realistic TWS noise. (3) Finally, CICA is assessed when applied on a long-term realistically synthesized SST data set to ensure its efficiency and applicability on different types of geophysical/climate time series.

2 Complex Independent Component Analysis (CICA)

Statistical analysis techniques aim at decomposing random variables, here stored in observed time series, into (empirical) modes or ‘sources’, which are estimated by assuming mutual orthogonality (Preisendorfer 1988) or mutual independence (Hyvärinen 1999a) between them. To generally formulate statistical decompositions, let us consider \mathbf{X} to be the data matrix that contains m sampled random variables (measured time series) with length of n . Here, \mathbf{X} contains $\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T, i = 1, \dots, p$ in its columns. We also assume that each column is temporally centered, i.e., the column-wise temporal means have already been removed. This assumption does not harm the general applicability and performance of statistical decomposition techniques as discussed in Cardoso (1999).

The widely used PCA is usually applied to extract few orthogonal modes from observations that represent the dominant part of their variance. Thus, by applying PCA on the data matrix \mathbf{X} , orthogonal modes can be estimated, which fairly well approximate the data as $\mathbf{X}_j = \bar{\mathbf{P}}_j \mathbf{\Lambda}_j \mathbf{E}_j^T$, where $j < \min(n, p)$ is the number of retained modes and \mathbf{X}_j is an approximation of \mathbf{X} . Each principal component (PC, \mathbf{P}), its associated singular value (stored indidiagonal entries of $\mathbf{\Lambda}$), and empirical orthogonal function (EOF, \mathbf{E}) represent an orthogonal mode of \mathbf{X} or its approximation \mathbf{X}_j . In practice, the \mathbf{j} modes are estimated by eigenvalue decomposition of the autocovariance matrix $\mathbf{C} = \frac{1}{n} \mathbf{X}^T \mathbf{X}$ (see, Forootan 2014, pp. 25–27). It is clear from the above definition that \mathbf{C} contains only the instantaneous time series, which justifies that PCA is a stationary approach since it does not consider any out-of-phase information about the time series in its criterion.

Forootan and Kusche (2012) follow (e.g., Comon 1994a, b; Aires et al. 2002) and formulate ICA as a rotated extension of the PCA transformation as

$$\mathbf{X} \simeq \mathbf{X}_j = \mathbf{P}_j \mathbf{R}_j \mathbf{R}_j^T \mathbf{E}_j^T, \quad (1)$$

where \mathbf{X} and \mathbf{X}_j are $n \times p$ temporally centered data matrices. The $n \times j$ and $p \times j$ matrices $\mathbf{P}_j (= \bar{\mathbf{P}}_j \mathbf{\Lambda}_j)$ and \mathbf{E}_j are derived from PCA (Preisendorfer 1988). To derive the ICA modes, an optimum $j \times j$ rotation matrix \mathbf{R}_j has to be defined that rotates either $\bar{\mathbf{P}}_j$ or \mathbf{E}_j while at

the same time making their columns as statistically independent as possible. Defining a proper rotation matrix (\mathbf{R}) requires an optimization of a measure of independence, for which one needs to use higher than two statistical moments (see different solutions of ICA in, e.g., Cardoso and Souloumiac 1993; Comon 1994a, b; Hyvärinen 1999a). This justifies that ICA is a higher-order statistical decomposition method. Considering Eq. (1), it is clear that ICA is a stationary approach since it relies on the same information retained by the PCA modes.

Complex ICA (CICA) is the focus of this paper and is derived in three steps: (Step 1) introducing non-stationary information by defining a new complex dataset; (Step 2) decomposing the complex data into orthogonal components using an eigenvalue decomposition technique; and (Step 3) rotating the orthogonal components to estimate independent patterns.

Step-1 In order to introduce non-stationary information into the decomposition procedure, we define a complex field (\mathbf{Y}) that contains the observed time series (\mathbf{X}) as its real part and their Hilbert transformation as its imaginary part (multiplied by $i = \sqrt{-1}$). Thus,

$$\mathbf{Y} = \mathbf{X} + i \mathcal{H}(\mathbf{X}), \quad (2)$$

where \mathcal{H} performs the Hilbert transformation and adds an approximation of the first time derivative of observations (\mathbf{X}) into Eq. (2), see the justification in Appendix 1.

Step-2 The complex field (\mathbf{Y}) in Eq. (2) is decomposed into orthogonal components using an eigenvalue decomposition as

$$\mathbf{Y} \simeq \mathbf{Y}_j = \bar{\mathbf{P}}_j^Y \mathbf{\Lambda}_j^Y \mathbf{E}_j^{YH}, \quad (3)$$

where the columns of the $n \times j$ matrix $\bar{\mathbf{P}}_j^Y$ are (temporal) complex principal components (CPCs). Similarly, the columns of the $p \times j$ matrix \mathbf{E}_j^Y contain complex values and correspond to the spatial components (CEOFs). Finally, H indicates the Hermitian transpose and the $j \times j$ matrix $\mathbf{\Lambda}_j^Y$ contains the real-number singular values in its main diagonal entries. The upper-index ‘Y’ is used to distinguish the decompositions derived from the new complex dataset.

The autocovariance ($\mathbf{C}^Y = \frac{1}{n} (\mathbf{X}^T \mathbf{X} + \mathcal{H}(\mathbf{X})^T \mathcal{H}(\mathbf{X}) + i(\mathbf{X}^T \mathcal{H}(\mathbf{X}) - \mathcal{H}(\mathbf{X})^T \mathbf{X}))$) used to perform the above decomposition (Eq. 3) contains information on the cross-spectral values, averaged over all frequencies ($-\pi < w_k < \pi$) that exist in the observed time series \mathbf{X} and their Hilbert transformation ($\mathcal{H}(\mathbf{X})$). Therefore, its decomposition yields complex orthogonal components of $\bar{\mathbf{P}}_j^Y$ and \mathbf{E}_j^Y , which retain the propagating disturbances present in the original data matrix \mathbf{X} . It should be mentioned here that, when a priori knowledge on the spectral frequency range of a certain pattern exists, then it is better to filter the original data and exclude those frequencies before implementing Eq. (3). This can be done by applying a band-limited filter (centered on the known frequency) to the data and its Hilbert transformation. Such pre-filtering will enhance extraction of not yet discovered cyclic or semi-cyclic patterns (Horel 1984).

Step-3 Complex independent components are estimated here by rotating the orthogonal components of Eq. (3), similar to the formulation in Eq. (1), as

$$\mathbf{Y} \simeq \mathbf{Y}_j = \mathbf{P}_j^Y \mathbf{R}_j^Y \mathbf{R}_j^{YH} \mathbf{E}_j^{YH}, \quad (4)$$

where here \mathbf{R}_j^Y is an $j \times j$ orthogonal rotation matrix ($\mathbf{R}_j^Y \mathbf{R}_j^{YH} = \mathbf{R}_j^{YH} \mathbf{R}_j^Y = \mathbf{I}_j$, and \mathbf{I} being the identity matrix), which should be defined.

Statistical independence is defined based on the probability density function (PDF) by stating that random variables are independent if and only if their joint distribution can be factorized to the product of their marginal distributions (Hyvärinen 1999b) as

$$p(\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_j) = \prod_{i=1}^j p(\mathbf{s}_i), \quad (5)$$

where \mathbf{s} is a discrete random variable taking values of s_1, s_2, \dots, s_n with probabilities of p_1, p_2, \dots, p_n , respectively. In Eq. (5), $p(\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_j)$ denotes the joint PDF and $p(\mathbf{s}_i)$ denotes the marginal PDF of each source. Therefore, in order to estimate independence, one needs to estimate either the joint and marginal PDFs of the rotated components ($\mathbf{P}_j^Y \mathbf{R}_j^Y$ or $\mathbf{E}_j^Y \mathbf{R}_j^Y$) in Eq. (4) or an approximation of their PDFs. In this study, the diagonalization of the fourth-order statistical cumulants (presented in Foorootan and Kusche 2012) is used as our approximation to find a proper \mathbf{R}_j^Y in Eq. (4).

Statistical properties of a random variable can be described by its statistical moments or, more conveniently, by its cumulants, denoted here as $K(x)$, which can be defined via the cumulant-generating function $g(t)$ as the logarithm of the moment-generating function $g(t) = \log[E(e^{tx})] = \sum_{n=1}^{\infty} \kappa_n \frac{t^n}{n!}$. Therefore, the cumulants κ_n can be obtained by n times differentiating the expansion of $g(t)$ and evaluating the result at zero, or $\kappa_n = \frac{\partial^n}{\partial t^n} g(t)|_{t=0}$. The cumulant of the sum of two statistically independent random variables s_1 and s_2 can be written as the sum of the cumulant of each, i.e., $\kappa_{s_1+s_2}(t) = \log[E(e^{t(s_1+s_2)})] = \log[E(e^{ts_1})E(e^{ts_2})] = \log[E(e^{ts_1})] + \log[E(e^{ts_2})] = \kappa_{s_1}(t) + \kappa_{s_2}(t)$. This can be simply extended to more than two random variables (see, e.g., Cardoso and Souloumiac 1993).

In general, the fourth-order cumulants of complex random variables can be defined as

$$\begin{aligned} K^Y(\mathbf{x}_i, \mathbf{x}_j^*, \mathbf{x}_k, \mathbf{x}_l^*) &= E(\mathbf{x}_i \mathbf{x}_j^* \mathbf{x}_k \mathbf{x}_l^*) - E(\mathbf{x}_i \mathbf{x}_j^*) E(\mathbf{x}_k \mathbf{x}_l^*) \\ &\quad - E(\mathbf{x}_i \mathbf{x}_k) E(\mathbf{x}_j^* \mathbf{x}_l^*) - E(\mathbf{x}_i \mathbf{x}_l^*) E(\mathbf{x}_j^* \mathbf{x}_k), \end{aligned} \quad (6)$$

where $E(\cdot)$ is the expectation operator, and $*$ represents the complex conjugate. Cardoso and Souloumiac (1995) show that for a multivariate case (several random variables), the fourth-order cumulant tensor can be defined using an arbitrary matrix \mathbf{M} as

$$\mathbf{Q}^Y(\mathbf{M}) = E((\mathbf{x}^* \mathbf{M} \mathbf{x})(\mathbf{x} \mathbf{x}^*)) - \mathbf{C}^Y \text{ trace}(\mathbf{M} \mathbf{C}^Y) - \mathbf{C}^Y(\mathbf{M} + \mathbf{M}^T) \mathbf{C}^Y, \quad (7)$$

with \mathbf{C}^Y being the autocovariance matrix. The required rotation matrix \mathbf{R}_j for diagonalizing Eq. (7) can be computed, for example, using the joint diagonalization (JD) approach as described below.

The cumulant tensor \mathbf{Q}^Y in Eq. (7) for either complex EOFs (CEOFs) or complex PCs (CPCs) contains n^4 entries, i.e., the number of fourth-order cross-cumulants. Using the JD algorithm, \mathbf{R}_j is found as the minimizer of the squared off-diagonal cumulant entries

$$f(\mathbf{R}_j^Y) = \sum_{t \neq s}^j f_{ts}^2 \quad \mathbf{F} = \sum_{m=1}^{n^2} \mathbf{R}_j^{YH} \mathbf{Q}^Y (\mathbf{M}_m) \mathbf{R}_j^Y, \quad (8)$$

where t and s represent the row and column of each entry. An optimum $j \times j$ orthogonal rotation matrix \mathbf{R}_j^Y ($\mathbf{R}_j^Y \mathbf{R}_j^{YH} = \mathbf{R}_j^{YH} \mathbf{R}_j^Y = \mathbf{I}_j$) is the solution of Eq. (8), which replaces the rotation matrix in Eq. (4) to identify complex independent components. It is worth mentioning that since the cross-cumulants in Eq. (6) or Eq. (7) are only an approximation of the statistical independence in Eq. (5), the optimization criterion in Eq. (8) can be solved with a restricted number of iterations. As a result, the statistical components derived from ICA/CICA might not always be ‘perfectly’ independent and rather ‘as independent as possible’ (Cardoso and Souloumiac 1995). For simplicity, however, here, we call the estimated components ‘independent.’

By choosing Eq. (8) to approximate independence, we assume that the joint distribution of de-correlated components is circular (see, e.g., Cardoso and Souloumiac 1995). This assumption does not harm the extraction of cyclic and semi-cyclic components. In order to separate sources with significantly non-circular joint distributions, one can use alternative independence criteria similar to, e.g., Fu et al. (2015). It is also worth mentioning that the strategy used to extend ICA to CICA shares the same view as in Horel (1984) for expending PCA/EOF to CEOF. In other words, we assume that, when the underlying processes are non-stationary, instantaneous observations and their out-of-phase records can be considered in decomposition; thus, the proposed CICA is formulated in the mentioned three steps. Other extensions of ICA to CICA exist that solve the signal separation problem in the frequency domain (e.g., Sawada et al. 2005). Anemüller et al. (2003, 2004) provide a frequency domain CICA algorithm and discuss its efficiency in extracting sources with limited spectral extents. Addressing pros and cons of different CICA techniques for analyzing geophysical and climate records requires further research.

Choice of Spatial Complex ICA (SCICA) and Temporal Complex ICA (TCICA): Forootan and Kusche (2012) indicate that ICA can be applied in two ways, from which (a) one can extract the statistically independent spatial patterns and their associated temporal patterns; the decomposition is known as spatial ICA (SICA), and (b) statistically independent temporal patterns and their associated spatial patterns are extracted, i.e., known as temporal ICA (TICA). In the SICA, observations are interpreted as a sequence of spatial snapshots; thus, one is interested in extracting stable spatial patterns from these observations. In the TICA, the hypothesis is that the sampled time series contain information of various time scales. Thus, searching for temporally independent patterns can extract distinguishable variability from them. Similar to a real-valued ICA, spatial complex ICA (SCICA method) and temporal complex ICA (TCICA) are derived from the following equations.

SCICA: Considering Eq. (4), one can define $\tilde{\mathbf{S}}^Y$ as statistically independent (complex) sources, and their associated (complex) temporal components can be derived as $\tilde{\mathbf{A}}^Y$.

$$\tilde{\mathbf{S}}_{p \times j}^Y = \mathbf{E}_j^Y \mathbf{R}_j^Y \quad \tilde{\mathbf{A}}_{n \times j} = \tilde{\mathbf{P}}_j^Y \mathbf{\Lambda}_j^Y \mathbf{R}_j^Y. \quad (9)$$

TCICA: The Hilbert transpose of Eq. (4) provides $\mathbf{Y}^H \simeq \mathbf{Y}_j^H = \mathbf{E}_j^Y \mathbf{\Lambda}_j^Y \mathbf{R}_j^Y \mathbf{R}_j^{YH} \mathbf{P}_j^{YH}$. Accordingly, the temporal components are derived as $\tilde{\mathbf{S}}$ and the corresponding spatial components are defined as $\tilde{\mathbf{A}}^Y$. Both components contain complex entries.

$$\mathbf{S}_{n \times j}^Y = \mathbf{P}_j^Y \mathbf{R}_j^Y \quad \mathbf{A}_{p \times j}^Y = \mathbf{E}_j^Y \mathbf{\Lambda}_j^Y \mathbf{R}_j^Y. \quad (10)$$

Details of implementing the SCICA and TCICA algorithms are described in Algorithm 1 of Appendix 2.

The temporal amplitudes and their associated phase patterns can be estimated from the complex independent components (CICs) in both SCICA and TCICA, which are formulated below, where the subindex ‘S’ represents spatial components and the subindex ‘T’ corresponds to the temporal components.

Spatial Amplitude: Amplitude_S in Eq. (11) of the l ’th component (with $l \in \{1, \dots, \min(n, p)\}$ and n and p to be the dimensions of \mathbf{X}) is derived as

$$\text{Amplitude}_S(\tilde{\mathbf{s}}_l) = (\tilde{\mathbf{s}}_l \bullet \tilde{\mathbf{s}}_l^*)^{1/2} \quad \text{and} \quad \text{Amplitude}_S(\mathbf{a}_l) = (\mathbf{a}_l \bullet \mathbf{a}_l^*)^{1/2}, \quad (11)$$

where $\tilde{\mathbf{s}}_l$ is a column of $\tilde{\mathbf{S}}^Y$ in SCICA and \mathbf{a}_l is a column of \mathbf{A}^Y in TCICA. Here, ‘ \bullet ’ is a vector valued operator that provides an element-by-element multiplication product of the columns.

Spatial Phase: Phase_S in Eq. (12), that corresponds to the l ’th component, is computed from

$$\text{Phase}_S(\tilde{\mathbf{s}}_l) = \arctan \left[\frac{\text{Im}(\tilde{\mathbf{s}}_l)}{\text{Re}(\tilde{\mathbf{s}}_l)} \right] \quad \text{and} \quad \text{Phase}_S(\mathbf{a}_l) = \arctan \left[\frac{\text{Im}(\mathbf{a}_l)}{\text{Re}(\mathbf{a}_l)} \right], \quad (12)$$

for SCICA and for TCICA, respectively.

Temporal Amplitude: Amplitude_T in Eq. (13) can be estimated as

$$\text{Amplitude}_T(\tilde{\mathbf{a}}_l) = (\tilde{\mathbf{a}}_l \bullet \tilde{\mathbf{a}}_l^*)^{1/2} \quad \text{and} \quad \text{Amplitude}_T(\mathbf{s}_l) = (\mathbf{s}_l \bullet \mathbf{s}_l^*)^{1/2}, \quad (13)$$

for SCICA and for TCICA, respectively.

Temporal Phase: Phase_T in Eq. (14) is estimated from

$$\text{Phase}_T(\tilde{\mathbf{a}}_l) = \arctan \left[\frac{\text{Im}(\tilde{\mathbf{a}}_l)}{\text{Re}(\tilde{\mathbf{a}}_l)} \right] \quad \text{and} \quad \text{Phase}_T(\mathbf{s}_l) = \arctan \left[\frac{\text{Im}(\mathbf{s}_l)}{\text{Re}(\mathbf{s}_l)} \right], \quad (14)$$

for SCICA and for TCICA, respectively. In order to estimate the phase patterns, the division in Eqs. (12) and (14) is computed element by element.

Independent Mode: the l ’th mode in Eq. (15) consists of a spatial and its associated temporal component. Each mode of variability derived from the CICA decomposition (including both TCICA and SCICA) can be estimated in a similar manner to the CEOF decomposition, i.e.,

$$\text{Mode}(\tilde{\mathbf{a}}_l) = \text{Re}(\mathbf{p}_l^Y \mathbf{r}_l^Y \mathbf{r}_l^{YH} \mathbf{e}_l^{YH}), \quad (15)$$

where $l < \min(n, p)$ and n and p are the dimensions of \mathbf{X} . Here, \mathbf{p}^Y , \mathbf{r}^Y and \mathbf{e}^Y represent a column of the matrices \mathbf{P}^Y , \mathbf{R}^Y , and \mathbf{E}^Y in Eq. (4), respectively. The operator $\text{Re}(\cdot)$ returns the real part of the reconstructed time series.

Signal Reconstruction: Equation (16) is applied to reproduce (an approximation of) the original matrix $\mathbf{X}(n \times p)$, after applying the CICA, one can use

$$\mathbf{X} \simeq \mathbf{X}_j = \text{Re}(\mathbf{P}_j^Y \mathbf{R}_j^Y \mathbf{R}_j^{YH} \mathbf{E}_j^{YH}). \quad (16)$$

In Appendix 1, we provide the necessary formulation for estimating a Hilbert transformation, while in Appendix 2, an algorithm is offered to implement CICA, and finally, the uncertainty of independent modes is discussed in Appendix 3.

3 Data

This sections describes the two sorts of data that are used for testing decomposition techniques and also for generating synthetic tests.

3.1 Global Terrestrial Water Storage (TWS)

To perform our investigations, we use monthly $1^\circ \times 1^\circ$ global terrestrial water storage (TWS) changes from the Gravity Recovery And Climate Experiment (GRACE, Tapley et al. 2004) satellite mission. Monthly TWS data along with their corresponding errors, which are based on the RL05 spherical harmonics from the Centre for Space Research (CSR) at the University of Texas, are downloaded from <https://grace.jpl.nasa.gov/data/get-data>. The data consist of 155 fields and cover 2002.29–2016.53. The effect of glacial isostatic adjustment is accounted for by applying the corrections from Wahr and Zhong (2013).

3.2 Global Sea Surface Temperature (SST)

Monthly reconstructed global $1^\circ \times 1^\circ$ Reynolds SST data (Reynolds et al. 2002) that are frequently used for climate studies are considered here as an example of a long-term (1982–2016) climate dataset. The data are downloaded from <https://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html>.

3.3 Southern Oscillation Index (SOI)

ENSO is a large-scale ocean–atmosphere interaction in the Tropical Pacific, which affects the climate of many regions of the Earth (Trenberth 1990; Forootan et al. 2016). El Niño refers to the negative phase on ENSO, and its opposite phase is known as La Niña. El Niño often produces dry years that cause drought, and the opposite happens during La Niña years. The SOI is downloaded from <https://www.ncdc.noaa.gov/teleconnections/enso/indicators/soi/>, which is a measure of the large-scale fluctuations in air pressure occurring between the western and eastern tropical Pacific (i.e., the state of the Southern Oscillation). Prolonged periods of negative (positive) SOI values coincide with abnormally warm (cold) ocean waters across the eastern tropical Pacific typical of El Niño (La Niña) episodes.

3.4 Pacific Decadal Oscillation (PDO) Index

PDO is often described as a long-lived ENSO pattern within the Pacific. The PDO's pattern is more stable than ENSO's, because its phase does not change sign for 20–30 years, while that of ENSO only lasts 6–18 months. Shifts in the PDO phase can have significant implications for global climate, affecting namely Pacific and Atlantic hurricane activity, droughts and floods. The PDO index is downloaded from <https://www.esrl.noaa.gov/psd/data/correlation/pdo.data>.

3.5 North Atlantic Oscillation (NAO) Index

NAO is the dominant pattern of atmospheric variability over the North Atlantic Ocean, especially in winter, which is usually measured as the difference in sea level pressure between Iceland and the Azores (Hurrell 2003). Since the Atlantic storms mostly affect climate in the Europe and USA, NAO has strong impact on the precipitation patterns in these regions. The NAO index is downloaded from <https://www.esrl.noaa.gov/psd/data/correlation/nao.data>.

3.6 Atlantic Multi-decadal Oscillation (AMO) Index

AMO represents long-duration changes in the sea surface temperature of the North Atlantic Ocean, with cool and warm phases that may last for 20–40 years and introduce SST difference of about 1–2°C between extremes. The warm phase of AMO prolongs droughts and vice versa during its cold phase. The AMO index is downloaded from <https://www.esrl.noaa.gov/psd/data/correlation/amon.us.data>.

4 Setting Up Three Synthetic Datasets

Extracting trend, acceleration, seasonality, and semi-cyclic behavior that are introduced by large-scale ocean–atmosphere interactions are important for climatic and geophysical interpretations. Therefore, we construct synthetic time series to investigate the performance of statistical decomposition techniques, including the introduced CICA, in separating these components from observations that contain a mixture of the mentioned components. In our experiments, we use two frequently used datasets of $1^\circ \times 1^\circ$ monthly global terrestrial water storage (TWS) time series from GRACE, as well as $1^\circ \times 1^\circ$ monthly sea surface temperature (SST) from Reynolds reanalysis (see Sect. 3 for more details). The length of GRACE observations is only 155 months, and that of SST is 420 months.

Assume $\mathbf{x}(\lambda, \phi, t)$ stores the time series of ‘real’ TWS or SST at a location with the longitude λ and the latitude ϕ with t being time in years. To generate synthetic data, we formulate the following regression:

$$\begin{aligned} \mathbf{x}(\lambda, \phi, t) = & \beta_0(\lambda, \phi) + \beta_1(\lambda, \phi) \cdot t + \beta_2(\lambda, \phi) \cos(2\pi t) + \beta_3(\lambda, \phi) \sin(2\pi t) \\ & + \beta_4(\lambda, \phi) \cos(4\pi t) + \beta_5(\lambda, \phi) \sin(4\pi t) + \beta_6(\lambda, \phi) S_1(t) \\ & + \beta_7(\lambda, \phi) \mathcal{H}(S_1(t)) + \beta_8(\lambda, \phi) S_2(t) + \beta_9(\lambda, \phi) \mathcal{H}(S_2(t)) + n(\lambda, \phi, t), \end{aligned} \quad (17)$$

in which the coefficients of the regression represent various time-scale variabilities of TWS and SST; β_0 as an offset depending on the start point of the time series, β_1 represents the linear trend, β_2 and β_3 for the annual cycle, β_4 and β_5 correspond to the semi-annual cycle, and finally $\beta_6, \beta_7, \beta_8$, and β_9 indicate changes due to ocean–atmosphere phenomena. A temporal error n is assumed to be normally distributed ($n \sim N(0, \sigma_n^2)$); therefore, it is independent from the base functions and stands for the deviations between observations and the fitted model. In Eq. (17), $S_1(t)$ and $S_2(t)$ can be introduced using climate indices introduced in Sect. 3. Temporal components used in Eq. (17) are of interest of various water resources and climate studies. For example, Fasullo et al. (2013), Eicker et al. (2016), and Forootan et al. (2016) applied this approach to analyze the impact of climate on global water storage and water fluxes, and precipitation over Australia, respectively. To account for the out-of-phase impact of these climate phenomena on TWS or SST, $\mathcal{H}(S_1(t))$

and $\mathcal{H}(S_2(t))$ are used that represent the time series of indices after being shifted by $\pi/2$ radian = 90° in the frequency domain [see Eqs. (19) and (20) in Appendix 1].

4.1 Accounting for ENSO While Generating Synthetic Global TWS Data

To introduce large-scale teleconnection impact on global TWS time series, in Eq. (17), we use the temporally normalized time series of the Southern Oscillation Index (SOI) and its Hilbert transform as $S_1(t)$ and $\mathcal{H}(S_1(t))$, respectively (both time series are shown in Appendix 1). In other words, to generate a synthetic TWS data, we account for ENSO's immediate and out-of-phase influence on TWS, with a fundamental assumption that the temporal behavior of TWS changes due to ENSO is similar to SOI. This might be in reality not perfectly true (see also Forootan et al. 2016), but the influence of this assumption is not strong enough to alter the results of our assessment, during which we try to measure the efficiency of decomposition techniques rather than focusing on a realistic estimation of climate impacts on geophysical data. We exclude $S_2(t)$ and $\mathcal{H}(S_2(t))$ while fitting Eq. (17) to TWS data since the length of data is limited and using other climate indices will inhibit a stable estimation.

4.2 Accounting for Dominant Teleconnections While Generating Synthetic SST Patterns Over the Pacific and Atlantic Oceans

We consider more complicated non-stationary patterns than in the previous section, while producing synthetic SST datasets. For this, 420 months of SST data from 1982–2016 is considered over the Atlantic Ocean (box of -66° – 13° E and -20° – 31° N), and the Pacific Ocean (box of 159° – 275° E and -30° – 19° N). Over the Atlantic Ocean, we fit Eq. (17) while replacing $S_1(t)$ with the normalized North Atlantic Oscillation (NAO) index and $S_2(t)$ with the normalized Atlantic Multi-decadal Oscillation (AMO) index. Over the Pacific Ocean, we use the same setup, but instead of Atlantic indices we use the normalized Southern Oscillation Index (SOI) for $S_1(t)$ and the normalized Pacific Decadal Oscillation (PDO) index for $S_2(t)$. The temperature data over the continents are not masked and included in our investigations. We should mention here that the spatial distribution, including extension and strength of input data, has an impact on the results of decomposition techniques as shown, for example, by Richman (1986). In this study, we do not focus on this issue and assume that the spatial extension of data is pre-defined. Our main aim here is to show how non-stationary information can be involved in the ICA procedure and its benefits are discussed.

Time series of the temporally normalized NAO and AMO are shown in Fig. 1 (top). Their spectral properties are estimated using the Least Squares Spectral Analysis (LSSA) technique, while the significance of the extracted frequencies is tested as in Sharifi et al. (2013), and the results are shown in Fig. 1 (bottom). Both time series are temporally distinguishable (their correlation coefficient is found to be 0.22, which is relatively low). Spectrally, the dominant frequencies of NAO and AMO are found to be well distinguishable, i.e., as expected, long-wavelength frequencies are dominant in the AMO and those of NAO are mainly related to seasonal to multi-year cycles (see Fig. 1, bottom).

Similarly, the time series of the normalized SOI and PDO are shown in Fig. 2 (top), and their power spectrum is shown in Fig. 2 (bottom). These two time series are found to be spectrally more similar than the Atlantic Indices, because many dominant frequencies are repeated in both time series. Their correlation coefficients are also 0.38, which is larger than that of NAO and AMO. Comparing power spectrum in Fig. 1 (bottom) and Fig. 2

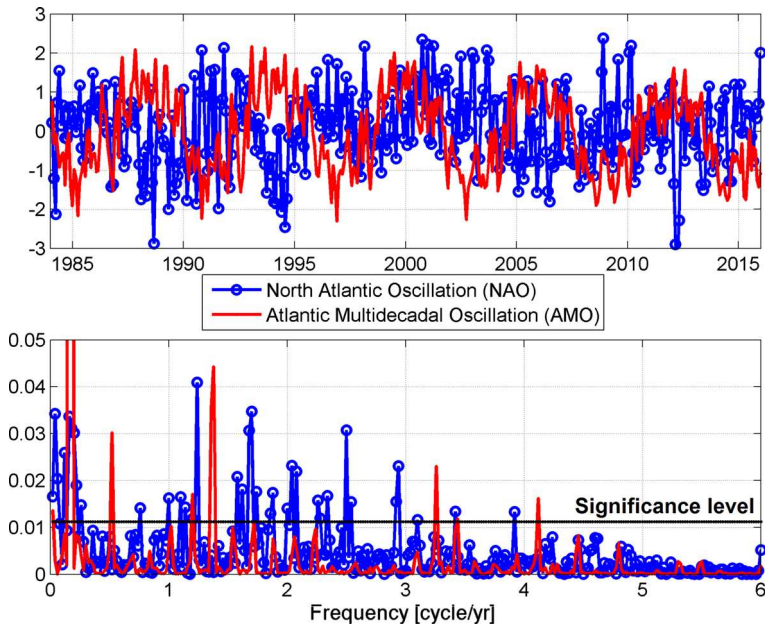


Fig. 1 The North Atlantic Oscillation (NAO) and the Atlantic Multi-decadal Oscillation (AMO) indices (top), and their associated power spectrums (bottom)

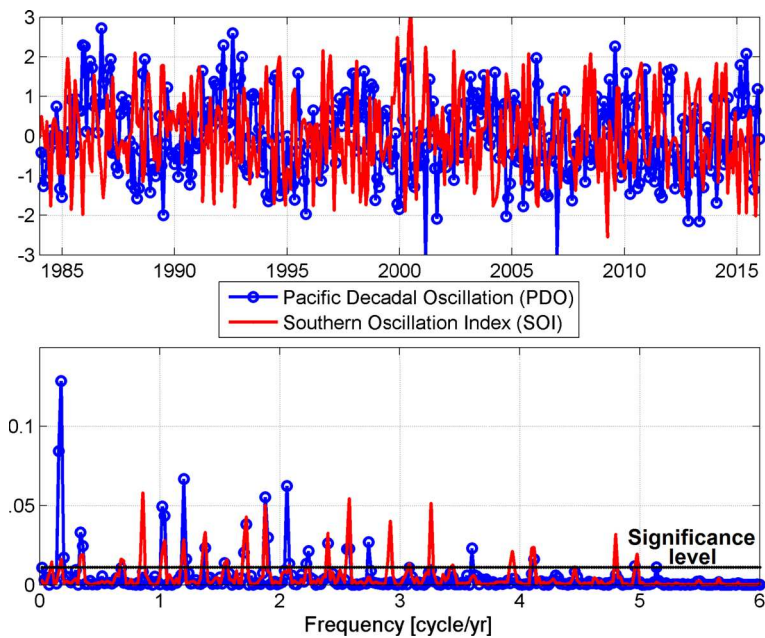


Fig. 2 The Southern Oscillation Index (SOI) and the Pacific Decadal Oscillation (PDO) indices (top), and their power spectrum (bottom)

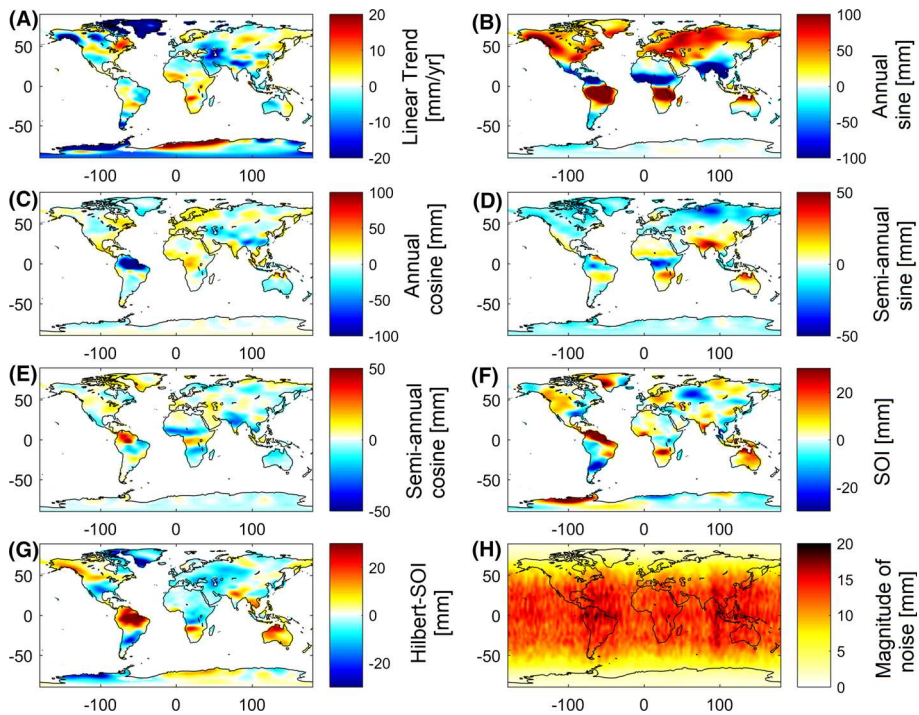


Fig. 3 Properties of the simulation used for assessing the decomposition techniques used in this study. Patterns are estimated by fitting Eq. (18) to 155 months of GRACE TWS covering 2002.29–2016.53: **a** Linear trend $\hat{\beta}_1$, **b** coefficient of the cosine part of the annual cycle $\hat{\beta}_2$, **c** coefficient of the sine part of the annual cycle $\hat{\beta}_3$, **d** coefficient of the cosine part of the semi-annual cycle $\hat{\beta}_4$, **e** coefficient of the sine part of the semi-annual cycle $\hat{\beta}_5$, **f** coefficient of the SOI $\hat{\beta}_6$, **g** coefficient of the out-of-phase SOI $\hat{\beta}_7$, and **h** standard deviation of the correlated noise

(bottom), we derive a correlation coefficient of 0.34 between NAO and AMO, while this is found to be 0.58 between the spectrum of SOI and PDO. Separating these similar patterns is difficult for decomposition techniques; therefore, the three synthetic setups should be well representative of real-world signal separation problems and a good testbed for the decomposition techniques assessed in this paper.

4.3 Constructing the Synthetic Data

By predefining the base functions in the form of Eq. (17), we have already determined the behavior of the synthetic time series. In the following section, we compute realistic amplitudes for them to be as close to the real datasets as possible. Therefore, regression coefficients of $\hat{\beta}_0 \dots \hat{\beta}_9$ are estimated using a least squares adjustment (LSA, Koch 1999). Synthetic global TWS and SST time series are then generated using these estimated coefficients as

$$\begin{aligned}\hat{\mathbf{x}}(\lambda, \phi, t) = & \hat{\beta}_0(\lambda, \phi) + \hat{\beta}_1(\lambda, \phi) \cdot t + \hat{\beta}_2(\lambda, \phi) \cos(2\pi t) + \hat{\beta}_3(\lambda, \phi) \sin(2\pi t) \\ & + \hat{\beta}_4(\lambda, \phi) \cos(4\pi t) + \hat{\beta}_5(\lambda, \phi) \sin(4\pi t) + \hat{\beta}_6(\lambda, \phi) S_1(t) \\ & + \hat{\beta}_7(\lambda, \phi) \mathcal{H}(S_1(t)) + \hat{\beta}_8(\lambda, \phi) S_1(t) + \hat{\beta}_9(\lambda, \phi) \mathcal{H}(S_2(t)) + \epsilon(\lambda, \phi, t).\end{aligned}\quad (18)$$

To illustrate this procedure, for example, the estimated coefficients that correspond to the linear trend, annual, semi-annual, and the ENSO-related variability in TWS data are shown in Fig. 3a–g. GRACE Tellus Web site provides TWS errors that are spatially correlated and temporally random. These error fields (see the standard deviation of the produced noise in Fig. 3h) are used as ϵ in Eq. (18). For generating synthetic SST data, we use spatially and temporally random fields with the standard deviation of 10% of the SST time series, which is higher than the amplitude of errors in the real SST data. Therefore, in each simulation, the true spatial and temporal solutions of the decomposition problem are known. Thus, one can compare the components extracted by statistical methods with the introduced patterns and evaluate the success of each method. In what follows, the introduced patterns of TWS are shown, and to avoid duplicated figures, only the extracted temporal components are presented. For the experiment with SST data, the introduced patterns are not shown, but instead both spatial and temporal components derived from decomposition techniques are discussed.

5 Assessing the Performance of Decomposition Techniques Using Global TWS Data

Before assessing the complex ICA (CICA), we first compare the two stationary methods of PCA and ordinary ICA when they are applied on the synthetic TWS data. Throughout this paper, our investigations are restricted to the ‘temporal’ version of ICA, i.e., TICA and TCICA, since we are interested in extracting temporally distinguished components as simulated in Eq. (18). Root mean squares of errors (RMSE) and correlation coefficients are used to measure similarity between the extracted components and the synthetic truth.

By definition, seven source signals (one linear trend, two annual, and two semi-annual cycles, as well as two ENSO-related sources) plus noise exist in the synthetic TWS data. Thus, we select the first seven dominant modes that are derived by applying PCA and TICA techniques. The first three modes, extracted by both techniques, are found to be very similar to the introduced linear trend (error of up to 3 mm/yr) and the annual sine and cosine cycles (error of up to 4 mm in amplitude, results are not shown here). This was somewhat to be expected as Forootan and Kusche (2013) indicate that ICA (Forootan and Kusche 2012) is able to perfectly separate an unknown mixture of trend and sinusoidal signals in the data, provided that the length of the dataset is infinite. Figure 4 shows the temporal patterns of the 4th to 7th modes, where those of PCA are mutually orthogonal and those of temporal ICA (TICA) are statistically as independent as possible. For brevity, the corresponding spatial components are not shown here.

By a simple visual comparison, one can clearly see that the TICA results are closer to the introduced signals. Particularly, TICA’s components in Fig. 4a, b are closer to the SOI and its Hilbert transformation (RMSE: 0.3) than those of PCA (RMSE: 0.7), and also its components in Fig. 4c, d better (than PCA) reproduce the introduced semi-annual cycles (RMSE of: 0.2 from ICA compared to 0.8 from PCA). We also observe that the spatial components of TICA are very similar to those in Figs. 3a–f. Another important result is that the semi-annual component is repeated in the four modes of PCA, and the high-amplitude ENSO peaks of 2010–2011 emerge in several orthogonal modes (clearly in PC4,

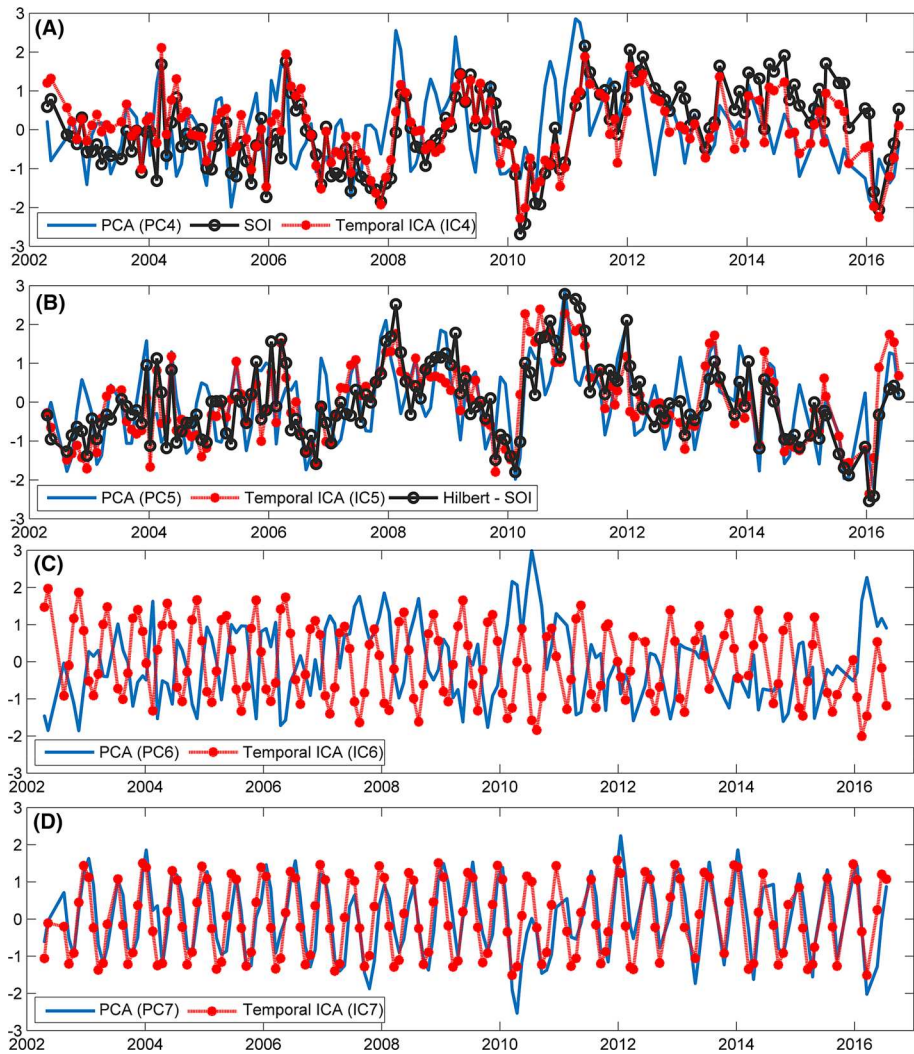


Fig. 4 Comparison of PCA and temporal ICA (TICA) for decomposing synthesized TWS data. Here only the temporal patterns of both techniques are shown, where PC4 and IC4 are shown in (a), PC5 and IC5 in (b), PC6 and IC6 in (c), and finally PC7 and IC7 in (d). We also show the ENSO index (SOI) and its Hilbert transformation in (b) and (a), respectively. All values are normalized; thus, they are unit-less

PC5, and PC6, Fig. 4a–c). Therefore, we confirm Forootan and Kusche (2012)’s previous conclusion that the PCA criterion, which seeks to retain the maximum variance in data, might not be adequate to isolate cyclic and semi-cyclic (here ENSO-like) sources.

It is worth mentioning here that the setup of the synthetic experiment might have influence on the performance of decomposition techniques. For example, we test another experiment by excluding the ENSO-related changes from Eq. (18). Then, PCA and TICA are applied on the new synthetic TWS data, for which the results indicate that both PCA and TICA techniques successfully extract the introduced linear trend, as well as annual and semi-annual cycles. The only difference is that a low amplitude annual cyclic is seen to be

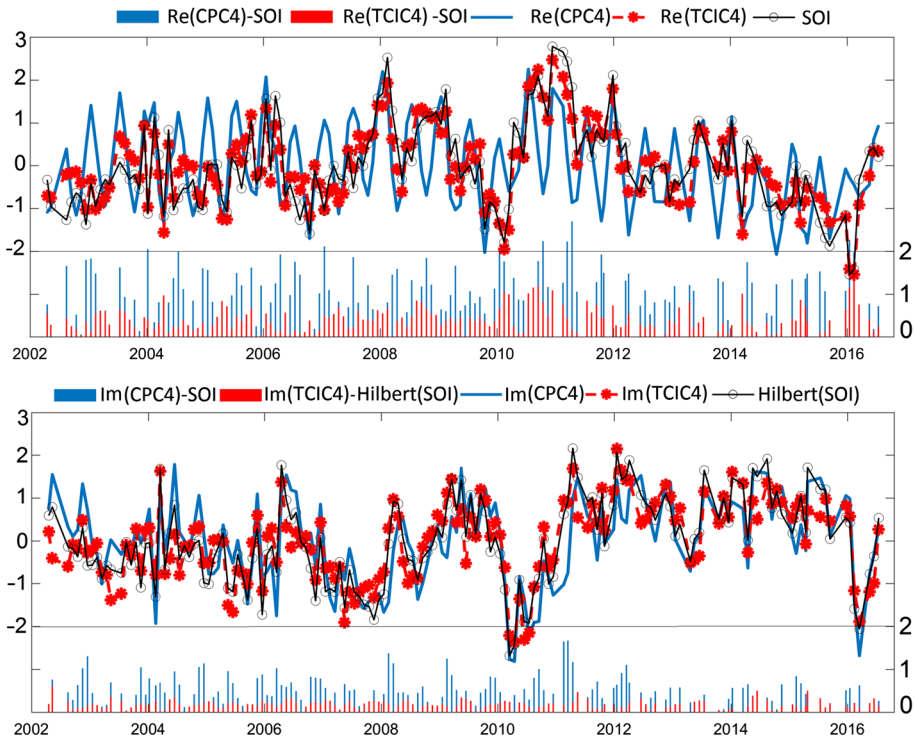


Fig. 5 Comparison of CEOF and temporal complex ICA (TCICA) for decomposing synthesized TWS data. (Top plot) shows the real part of the complex PC4 (CPC4) and temporal complex IC4 (TCIC4), where (bottom plot) shows their imagery parts. We also show the ENSO index (SOI) and its Hilbert transformation on top and bottom, respectively. All values are normalized and are thus unit-less. Blue and red bars indicate the absolute differences between the ENSO index or its Hilbert transformed pattern and CPC4 or TCIC4

mixed with the linear trend extracted by PCA (results are not shown). Based on these experiments, we conclude that the simulated ENSO-related patterns contain cycles that are close to the semi-annual cycle. Furthermore, its globally averaged standard deviation is comparable with that of the semi-annual cycle. As a result, both ENSO-related and semi-annual patterns are repeated in several modes as shown in Fig. 4.

We continue our investigation by testing the non-stationary techniques of CEOF (Eq. 3) and temporal complex ICA (TCICA, Eq. 10) on the synthetic TWS data of Eq. (18) including ENSO signal. As expected, CEOF and TCICA are able to separate the linear trend (mode 1), the annual (mode 2), semi-annual (mode 3), and ENSO components (mode 4) in their first four modes. Thus, unlike the stationary PCA and ICA techniques, they do not need seven modes. In Fig. 5, we compare the temporal components of the fourth mode (complex PC4, i.e., CPC4, and temporal complex IC4, i.e., TCIC4), where they likely correspond to the ENSO index (SOI). In Fig. 5 (top) and (bottom), the (absolute) differences of them with SOI are shown, from which the results indicate that the CICA results are closer to the simulation (RMSE: 0.4 from CICA and 0.6 from CEOF). The results in Fig. 5 also indicate that, similar to PCA, CEOF also suffers from the mixing problem, meaning that individual spatial and temporal modes of similar variances repeat in a number of orthogonal modes. For example, the semi-annual term can be clearly seen in the real part of CPC4 (Fig. 5, top).

It is worth mentioning that by multiplying the differences in Fig. 5 (top) and (bottom) with the corresponding spatial patterns (Fig. 3f, g), one can estimate TCICA error in

extracting ENSO from global TWS data, for which we estimate errors of up to 0.5–1 cm in terms of equivalent water height. These results therefore indicate that the current level of noise in the filtered TWS data does not have a dramatic impact on the performance of statistical techniques in extracting (semi-)cyclic behaviors with the period of longer than one year (see similar discussions in Kusche et al. 2016). Possible error sources occurred while decomposing global TWS data are investigated in Talpe et al. (2017).

6 Assessing the Performance of Decomposition Techniques Using Long-Term SST Data

In the previous section, we showed how introducing the higher-order statistical moments in the form of ICA and adding non-stationary information in the form of CICA can improve separation of ENSO from cyclic (seasonal) signals, although the length of observations was only 155 months. In this section, we test whether CICA can (a) separate semi-cyclic patterns that are spectrally similar, and (b) be successfully applied to any types of geophysical or climate time series. Therefore, here, long-term sea surface temperature (SST) data are used to test statistical decomposition techniques. In the light of results in previous section, the discussions are restricted to the comparisons between CEOF and TCICA. Furthermore, only the teleconnection patterns are compared below, instead of assessing the trend and annual and semi-annual cycles.

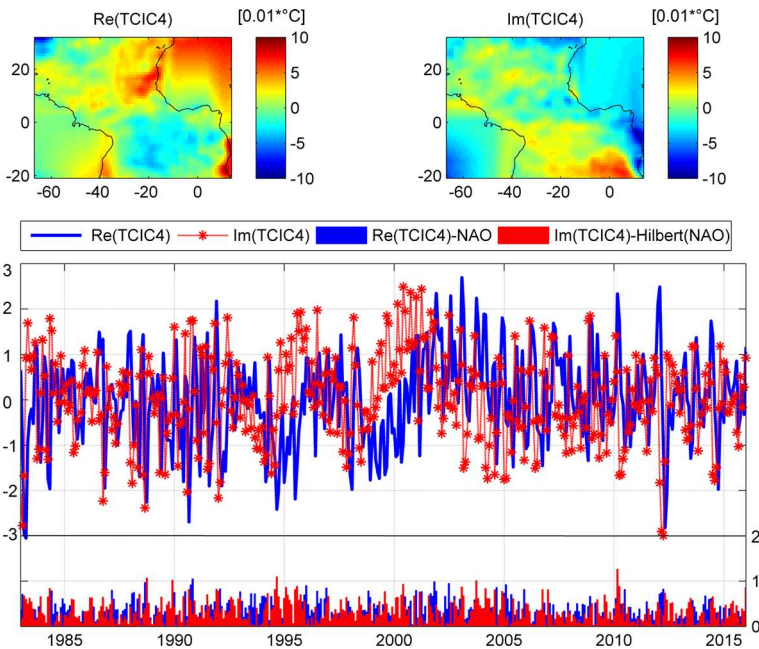


Fig. 6 Performance of the temporal complex ICA (TCICA) in extracting NAO from SST data. (Top) Spatial patterns of the real (left) and imaginary (right) parts of the fourth independent mode. Both anomaly maps have a unit of $^{\circ}\text{C}$ percent. (Bottom) Real and imaginary parts of the fourth normalized temporally independent component are shown. Blue and red bars indicate the absolute differences between the temporal components and NAO or its Hilbert transformed time series

6.1 Comparisons Over the Atlantic Ocean

The fourth and fifth independent modes from TCICA over the Atlantic Ocean are shown in Figs. 6 and 7, which correspond to the SST changes due to the North Atlantic Oscillation (NAO) and the Atlantic Multi-decadal Oscillation (AMO). Two spatial maps on the top of both figures indicate the real and imaginary parts of the spatial component and the corresponding temporal patterns are shown in the bottom of both figures. In both figures, we show the absolute values of the differences between the TCICA-derived temporal components and the NAO or AMO indices as well as their Hilbert transformed time series. The results indicate that the maximum value of the absolute differences is rarely greater than 1 (mean of 0.4), which corresponds to less than $\sim 16\%$ of NAO or AMO signal. This can be estimated using Eq. (16) while inserting the components in Figs. 6 and 7.

To compare with the TCICA results, the temporal components of the fourth and fifth mode of CEOF over the Atlantic Ocean are shown in Fig. 8. The results indicate that the multi-decadal pattern of AMO is repeated in both modes. This makes the maximum value of absolute differences considerably larger than TCICA, i.e., in some instances 3 (see the error bars in Fig. 8, bottom). In other words, the magnitude of errors in extracting NAO or AMO using CEOF (mean of 0.9) reaches up to $\sim 50\%$ of the signal itself.

In Fig. 6, the SST pattern preceding the NAO extends through the ocean from the northeastern side of Atlantic (from $\sim 20^\circ\text{N}$ see the plot of top left), extending down to $\sim 20^\circ\text{S}$ (plot on the top right), so that the same sign is seen north and south of the equator. One can see this propagation by estimating the spatial phase map using Eq. (12), where $\tilde{\mathbf{a}}$

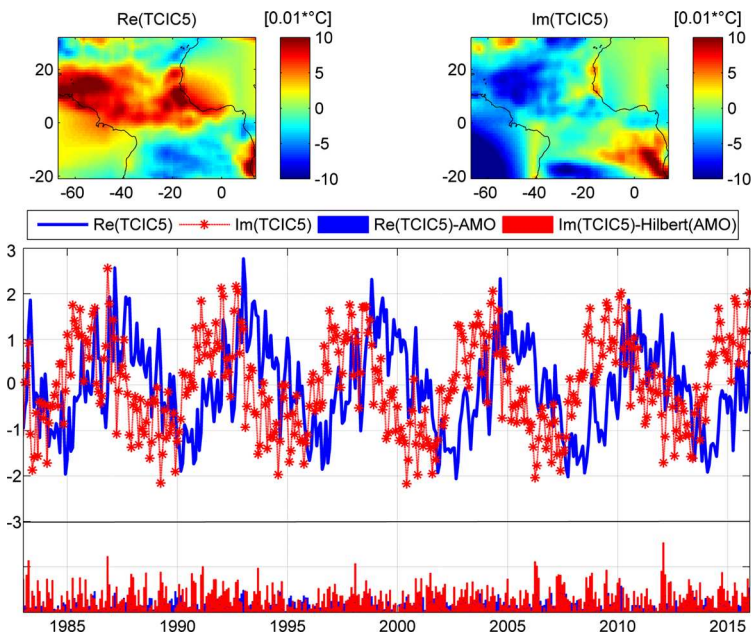


Fig. 7 Performance of the temporal complex ICA (TCICA) in extracting AMO from SST data. (Top) Spatial patterns of the real (left) and imaginary (right) parts of the fifth independent mode in degrees centigrade ($^\circ\text{C}$) percent. (Bottom) Real and imaginary parts of the fifth temporally independent component are shown that both are normalized. Blue and red bars indicate the temporal components and the AMO index or its Hilbert transformed time series

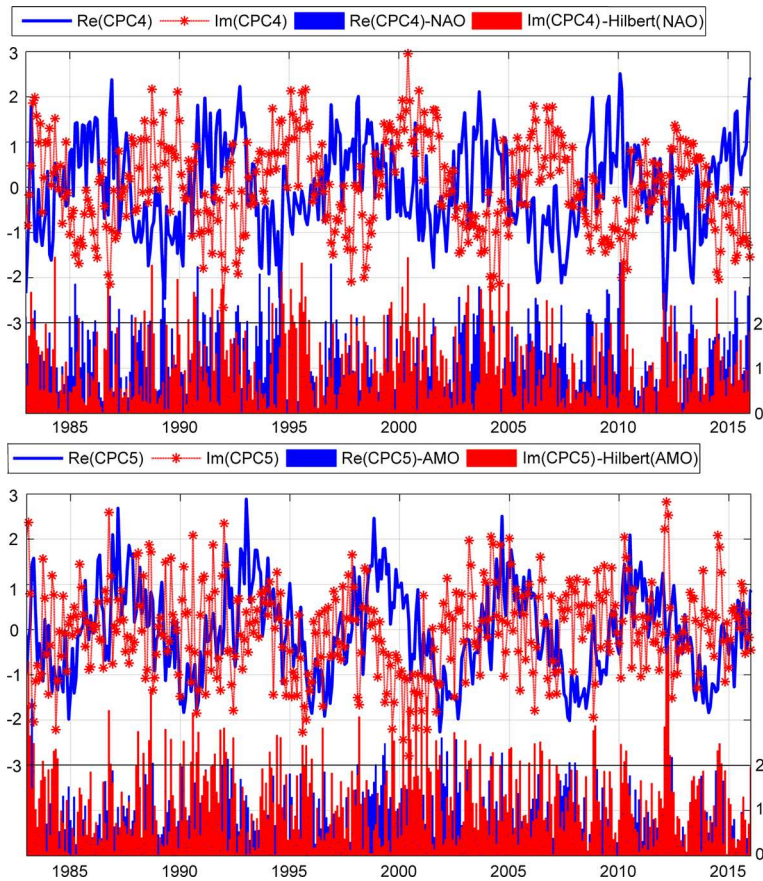


Fig. 8 Performance of CEOF in extracting NAO and AMO from SST data. (Top) real and imaginary parts of the fourth orthogonal mode. (Bottom) Real and imaginary part of the fifth normalized temporal component is shown. Blue and red bars indicate the absolute differences between the temporal components and NAO or AMO, as well as their Hilbert transformed time series

contains the spatial maps shown on the top plots in Fig. 6. The results are, however, not shown here. AMO is found to be spatially distributed as a dipole pattern with the equator in the middle (see Fig. 7, top left and top right plots). Considering the temporal patterns in Fig. 7 (bottom), one could expect cold or warm persistence SST in the Atlantic with the periods of ~ 25 years.

6.2 Comparisons Over the Pacific Ocean

The fourth and fifth independent modes from TCICA applied on the SST data within the Pacific Ocean are shown in Figs. 9 and 10, respectively. Note that the first three modes are related to the trend, annual, and semi-annual cycles that are not shown here. Both figures indicate that the spatiotemporal changes due to the ENSO and the Pacific Decadal Oscillation (PDO) are fairly well extracted in two separate independent modes. However, since the two phenomena are spatially and temporally correlated (see the discussions in

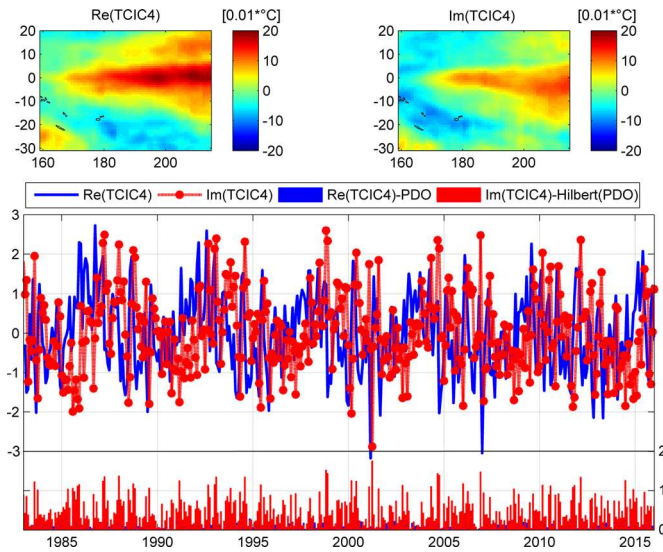


Fig. 9 Performance of the temporal complex ICA (TCICA) in extracting PDO from SST data. (Top) Spatial patterns of the real (left) and imaginary (right) parts of the fourth independent mode. Both anomaly maps have a unit of $^{\circ}\text{C}$ percent. (Bottom) The real and imaginary part of the fourth normalized temporally independent component. Blue and red bars indicate the absolute differences between the temporal components and PDO or its Hilbert transformed time series

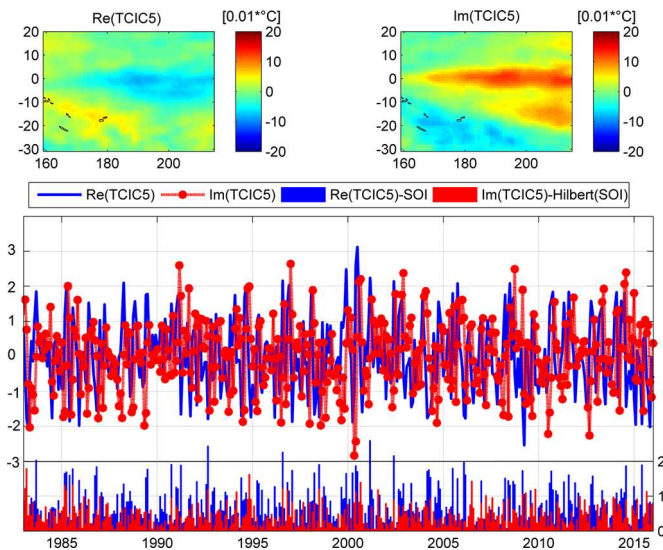


Fig. 10 Performance of the temporal complex ICA (TCICA) in extracting SOI from SST data. (Top) Spatial patterns of the real (left) and imaginary (right) parts of the fifth independent mode in $^{\circ}\text{C}$ percent. (Bottom) The real and imaginary part of the fifth temporally independent component that is normalized and is thus unit-less. Blue and red bars indicate the absolute differences between the temporal components and the SOI index or its Hilbert transformed time series

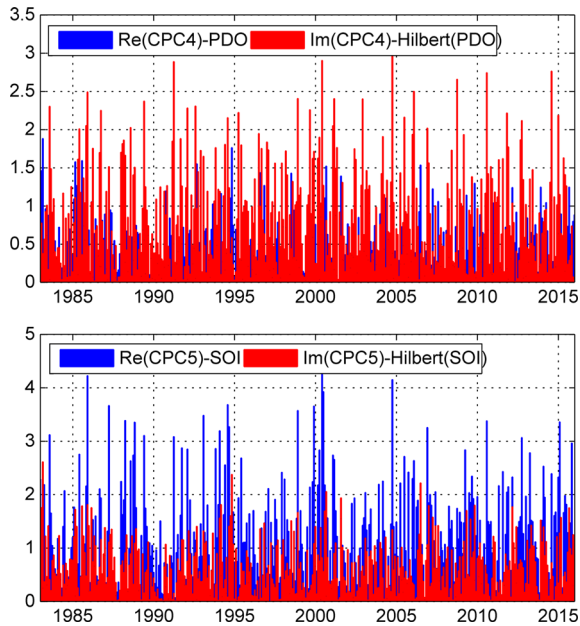


Fig. 11 Performance of CEOF in extracting SOI and PDO from SST data over the Pacific Ocean. Blue and red bars indicate the absolute differences between the temporal components of the fourth orthogonal mode and PDO as well as its Hilbert transformed time series (Top). Similar differences, to the top, but between the temporal components of the fifth orthogonal mode and SOI as well as its Hilbert transformed time series (Bottom). Time series are unit-less

Sect. 4.2), the absolute values of the differences between the TCICA-derived temporal components and the PDO or SOI indices as well as their Hilbert transformed time series reach up to 2 (mean of 0.7, see the bar plot on the bottom of Figs. 9 and 10). These differences are bigger than those over the Atlantic Ocean, and we expect that the overall error that corresponds to the separation of SOI from PDO reaches up to $\sim 30\%$ of the signal. The performance of CEOF in separating the SOI- and PDO-related patterns is found to be much worse than TCICA. In Fig. 11, the absolute differences between the temporal components of the fourth and fifth orthogonal modes of CEOF over the Pacific Ocean and SOI or PDO as well as their Hilbert transformed time series are shown. In both graphs, evident cyclic behavior can be seen, whose amplitudes reach to the amplitude of the SOI or PDO signals (mean of 1.2). It is worth mentioning that the considered level of SST noise (10% of signal) does not significantly harm the decomposition results.

7 Summary and Conclusions

In recent decades, decomposition techniques have garnered increasing interest for analyzing geophysical time series. In this study, we discussed the mathematical details of a number of frequently used statistical decomposition techniques, namely principal component analysis (PCA)/empirical orthogonal function (EOF), the more recent independent component analysis (ICA), and complex EOF (CEOF). With these existing techniques in mind, a novel decomposition technique, called complex ICA (CICA), is introduced. CICA combines the

advantage of an ordinary ICA (Forootan and Kusche 2012), i.e., involving higher (than two)-order statistical information embedded in the data into the decomposition procedure, and non-stationary information as in CEOF (Horel 1984).

The mathematical details of CICA are described in detail, and an algorithm to implement the method has been provided (see Sect. 2 and Appendix 2). Three synthetic datasets are also generated to test the proposed CICA technique in separating climate-related patterns from multivariate terrestrial water storage (TWS, 2003–2016) and sea surface temperature (SST, 1982–2016) time series. Our results indicate that CICA considerably mitigates the clustering behaviors that usually occur after application of the second-order statistical decomposition techniques. CICA also captures stationary and non-stationary variability in both TWS and SST data in fewer number of modes. Particularly, we show that, given the time series to be long enough (e.g., SST data used here), CICA can separate complicated semi-cyclic patterns such as those of the El Niño Southern Oscillation (ENSO) from the Pacific Decadal Oscillation (PDO), and the North Atlantic Oscillation (NAO) from the Atlantic Multi-decadal Oscillation (AMO).

The orthogonal property of PCA and CEOF decomposition is very useful since the covariance matrix of any subset of retained modes is always diagonal. Both techniques also capture the dominant part of the variance in the original dataset; therefore, their application for dimension reduction is recommended. However, when the PCA or CEOF components are treated individually, their results can be misleading since they mix physical processes with similar variance properties (see Figs. 8, 11). In those cases, ICA and CICA are shown to be better suited. Computational complexity of ICA and CICA is, however, higher than of second-order techniques. Therefore, for those applications that require one to extract a portion of total variance, for example, in dimension reduction studies, rather than interpreting individual modes, second-order statistical techniques (PCA/CEO) might be a better choice. CEOF and CICA are found to be more efficient than PCA and ICA when the input data contain non-stationary information. For example, using complex techniques, smaller number of modes requires to retain a certain portion of the total variance in the original data.

A reliable estimation of sample length and uncertainty estimation of the CICA derived modes is discussed in Appendix 3. Our numerical investigations indicate that a minimum length of 100 months is required to separate linear trend, annual, and semi-annual cycles, as well as the semi-cyclic ENSO from GRACE TWS data in the presence of realistic noise. The assessment, however, only considers the statistical and numerical errors in estimating statistically independent components, and the minimum length that is required to accurately represent all spectral properties of the ENSO index has not been considered.

The ICA criterion applied in this study is based on the joint diagonalization of the fourth-order cumulants, which has been generalized by, e.g., Moreau (2001) to include a variety of higher-order cumulants. In another attempt, Fu et al. (2015) provide a CICA algorithm that exploits three types of statistical properties, i.e., non-Gaussianity, non-whiteness, and non-circularity, to ensure the best possible approximation of statistical independence. Such extensions can be applied to improve the estimation of independence when the time series are long enough, such as SST data in this study. Applying ICA/CICA that requires computing more statistical moments from the length-limited time series, such as those of GRACE TWS, might itself introduce unwanted uncertainties. A rigorous investigation of such extensions will be addressed in future research. The CICA technique will be applied in future contribution to extract new ENSO indices from ‘real’ datasets such as GRACE TWS, SST, and global precipitation.

Acknowledgements The authors are grateful to Professor M. Rycroft (Editor in Chief) and two anonymous reviewers for providing helpful comments, which are used to improve the quality of this manuscript. We are

also grateful to the GRACE and sea surface temperature (SST) data, as well as climate indices that are freely available from providers. This research is partially supported by the Belmont Forum/IGFA G8 Coastal Vulnerability project, Band-Aid, <http://Belmont-BandAid.org>. The US component of the project is supported by the US National Science Foundation (Grant No. CER-1342644).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1: Hilbert Transformation

Assume the columns of the data matrix \mathbf{X} (shown by \mathbf{x}_t) to be a scalar field. This time series can be represented by a discrete Fourier transformation as

$$\mathbf{x}_t = \sum_k \mathbf{a}(\omega_k) \cos(\omega_k t) + \mathbf{b}(\omega_k) \sin(\omega_k t), \quad t = 1, \dots, n, \quad (19)$$

where $\mathbf{a}(\omega_k)$ and $\mathbf{b}(\omega_k)$ are vector Fourier coefficients at frequency ω_k (with accepting values between $-\pi$ and π that are selected according to the sampling rate). Thus, values of ω_k have to satisfy the Nyquist frequency rule and its equivalent wavelength should not exceed the length of time series (Chatfield 1989, chapter 7).

The Hilbert Transformation introduces a $\pi/2$ phase shift in the frequency domain to each individual time series (\mathbf{x}_t). Thus, the Hilbert transform ($\mathcal{H}(\cdot)$) of \mathbf{x}_t is given by $\mathcal{H}(\mathbf{x}_t) = \sum_{\omega_k} \mathbf{a}(\omega_k) \cos(\omega_k t + \pi/2) + \mathbf{b}(\omega_k) \sin(\omega_k t + \pi/2)$, $t = 1, \dots, n$, or

$$\mathcal{H}(\mathbf{x}_t) = \sum_k \mathbf{b}(\omega_k) \cos(\omega_k t) - \mathbf{a}(\omega_k) \sin(\omega_k t), \quad t = 1, \dots, n. \quad (20)$$

Comparing Eq. (20) with (19), it is clear that the resulted expansion in Eq. (20) is similar to the first derivative of Eq. (19) with a difference that the frequency argument (ω_k) is not involved in the Hilbert Transformation. For example, Fig. 12 shows the time series of SOI and its Hilbert transformed time series.

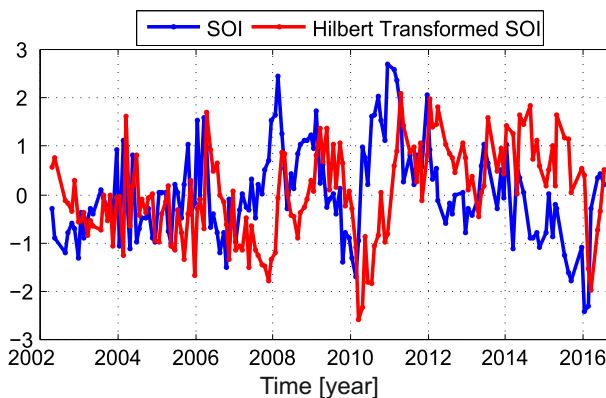


Fig. 12 Time series of normalized monthly SOI and its Hilbert transformed time series

Appendix 2: Spatial and Temporal Complex ICA Algorithms

Here, we present the CICA algorithm for both SCICA and TCICA, see Algorithm 1. The complex field \mathbf{Y} is generated by adding the Hilbert transform of Eq. (3) as the imaginary part to the observed data matrix \mathbf{X} , i.e., $\mathbf{Y} = \mathbf{X} + i \mathcal{H}(\mathbf{X})$. This is reflected in the stage 1. In stage 2, \mathbf{Y} is de-correlated by applying an eigenvalue decomposition, which provides the complex components of $\bar{\mathbf{P}}^Y$ and \mathbf{E}^Y . In stage 3, one must select one of the transformations $\mathbf{Z}_{\text{Spatial}} = \hat{\mathbf{W}}_{\text{Spatial}} \mathbf{Y}$ or $\mathbf{Z}_{\text{Temporal}} = \hat{\mathbf{W}}_{\text{Temporal}} \mathbf{Y}^H$. In stage 4, the fourth-order cumulant tensor \mathbf{Q} is built using Eq. (7), while considering the entries of \mathbf{Z} as its inputs, and in stage 5, an orthogonal rotation $\hat{\mathbf{R}}^Y$ is computed using the JD algorithm (Eq. 8). In stage 6, a mixing matrix is derived, which is used to estimate spatially or temporally independent components ($\hat{\mathbf{S}}_{\text{SCICA}}$ or $\hat{\mathbf{S}}_{\text{TCICA}}$) in stage 6.

Algorithm 1: Complex ICA algorithm based on diagonalization of the fourth order cumulants

Data: \mathbf{X}

Result: complex source signals $\mathbf{S}^Y/\tilde{\mathbf{S}}^Y$, and corresponding spatial/temporal components $\mathbf{A}^Y/\tilde{\mathbf{A}}^Y$

Default: $\xi = 10^{-2}/\sqrt{(n)}$, n is the length of time series;

1: Add the Hilbert transformation of \mathbf{X} to the original data and derive the complex data \mathbf{Y} Eq. (2);

2: Estimate a whitening matrix $\hat{\mathbf{W}}$ as $\hat{\mathbf{W}}_{\text{Spatial}} = \mathbf{\Lambda}^{-1}\bar{\mathbf{P}}^H$ or $\hat{\mathbf{W}}_{\text{Temporal}} = \mathbf{\Lambda}^{-1}\mathbf{E}^H$;

3: Compute de-correlated components by setting $\mathbf{Z}_{\text{Spatial}} = \hat{\mathbf{W}}_{\text{Spatial}}\mathbf{Y}$ or $\mathbf{Z}_{\text{Temporal}} = \hat{\mathbf{W}}_{\text{Temporal}}\mathbf{Y}^H$;

4: Estimate the cumulant tensor \mathbf{Q}_Z using the entries of \mathbf{Z} from Eq. (7) and order it in Eq. (8);

5: Find a rotation matrix $\hat{\mathbf{R}} = \prod_i \hat{\mathbf{V}}_i$ as the minimizer of the squared off-diagonal cumulant entries of Eq. (8);

while Sweep each cumulant matrix of Eq. (8) and derive $\hat{\mathbf{V}}_i(\theta)$ **do**

if off-diagonal elements are zero **then**

Select another pair;

else

θ is smaller than $\theta_{\min} = \xi$;

Select another pair;

end

end

6: Estimate a mixing matrix for SCICA as $\hat{\mathbf{W}}_{\text{SCICA}} = \hat{\mathbf{R}}^H \hat{\mathbf{W}}_{\text{Spatial}}$ or for TCICA as $\hat{\mathbf{W}}_{\text{TCICA}} = \hat{\mathbf{R}}^H \hat{\mathbf{W}}_{\text{Temporal}}$;

7: Estimate the source signals of $\tilde{\mathbf{S}}_{\text{SCICA}}$ as $\hat{\mathbf{W}}_{\text{SCICA}}\mathbf{Y}$ or $\mathbf{S}_{\text{TCICA}}$ as $\hat{\mathbf{W}}_{\text{TCICA}}\mathbf{Y}^H$. The corresponding patterns are derived as

$\tilde{\mathbf{A}}_{\text{SCICA}} = \bar{\mathbf{P}}_j^Y \mathbf{\Lambda}_j^Y \mathbf{R}_j^Y$ and $\hat{\mathbf{A}}_{\text{TCICA}}^Y = \mathbf{E}_j^Y \mathbf{\Lambda}_j^Y \mathbf{R}_j^Y$

Appendix 3: Reliability and Uncertainty of CICA Modes

The proposed ICA/CICA methods (Eqs. 1, 4) consist of a de-correlation step followed by a rotation (Eqs. 9, 10). Both steps are based on statistical estimates (Eq. 7) and solving a numerical optimization (Eq. 8). Given a known fourth-order cumulant tensor \mathbf{Q} , one needs to evaluate the minimum sample size N necessary to obtain a reliable estimation of tensor $\hat{\mathbf{Q}}_N$, which is close enough to the original \mathbf{Q} . Assuming that $\hat{\mathbf{Q}}_N$ and \mathbf{Q} are symmetric and positive definite, a distance measure can be defined as

$$d_N^2(\hat{\mathbf{Q}}_N, \mathbf{Q}) = \text{trace} \left(\log^2 \left(\hat{\mathbf{Q}}_N^{-1/2} \mathbf{Q} \hat{\mathbf{Q}}_N^{-1/2} \right) \right). \quad (21)$$

In an ideal case, each section of $\hat{\mathbf{Q}}_N^{-1/2} \mathbf{Q} \hat{\mathbf{Q}}_N^{-1/2}$ is equal to an identity \mathbf{I}_N , and consequently, d becomes 0. In real cases, Eq. (21) evaluates as

$$d_N^2(\hat{\mathbf{Q}}_N, \mathbf{Q}) = \text{trace}(\log^2(\mathbf{I}_N + \epsilon)) = \sum_{i=1}^N \log^2(1 + \epsilon_i) \approx N\epsilon^2, \quad (22)$$

where ϵ_i is a symmetric error matrix. In Eq. (22), we assume that these errors are equally distributed in each section of \mathbf{Q} . This equation suggests that the separation error is related to the number of independent of modes, i.e., reflected in N , and length of observations, which changes the magnitude of ϵ_i .

An exponential function $\exp(\cdot)$ can be used to represent the likelihood of errors in Eq. (22) as

$$\rho(d_N) = \exp \left(-1/2 \cdot d_N^2(\hat{\mathbf{Q}}_N, \mathbf{Q}) \right), \quad (23)$$

where $\rho(d_N)$ varies between 0 and 1, with 1 being the perfect match of $\hat{\mathbf{Q}}_N$ and \mathbf{Q} . one may assume that $\rho(d_N)$ should be larger than 0.95 in order to be representative of an acceptable estimation (see also Wu et al. 2008).

To numerically test this criterion, we consider the fourth cumulant tensor derived from a synthetic GRACE TWS example with a length of 155 months as our truth. Then, in each step, 5 months is randomly excluded from the synthetic dataset and Eq. (23) is evaluated. In all estimations, seven independent modes (sources) are extracted similarly to the original example shown in Sect. 4.1. This experiment indicates that a sample length of longer than 100 months (without sequential data gaps) is required to guarantee a $\rho(d_N)$ that is larger than 0.95.

Uncertainty Estimation of ICA Modes

In order to estimate the uncertainty of independent modes, one can follow a Bootstrap approach (Efron 1979) by generating m realizations of the data matrix $\mathbf{X}^i, i = 1, \dots, m$ by considering the original data \mathbf{X} and adding errors from its covariance matrix. Applying Eq. (4) on these realizations yields m sets of components, from which the uncertainty of independent modes can be estimated. Discussions on the uncertainty of decomposition techniques applied on GRACE TWS data are demonstrated in Talpe et al. (2017).

References

- Aires F, Rossow WB, Chedin A (2002) Rotation of EOFs by the independent component analysis: toward a solution of the mixing problem in the decomposition of geophysical time series. *J Atmos Sci* 59:111–123. [https://doi.org/10.1175/1520-0469\(2002\)059<0111:ROEBTI>2.0.CO;2](https://doi.org/10.1175/1520-0469(2002)059<0111:ROEBTI>2.0.CO;2)
- Anemüller J, Sejnowski TJ, Makeig S (2003) Complex independent component analysis of frequency-domain electroencephalographic data. *Neural Netw* 16:1311–1323. <https://doi.org/10.1016/j.neunet.2003.08.003>
- Anemüller J, Sejnowski TJ, Makeig S (2004) Reliable measurement of cortical flow patterns using complex independent component analysis of electroencephalographic signals. In: Puntonet CG, Prieto A (eds) *ICA 2004, LNCS 3195*. Springer, Berlin, Heidelberg, pp 1009–1016
- Awange J, Forootan E, Kuhn M, Kusche J, Heck B (2014) Water storage changes and climate variability within the Nile Basin between 2002–2011. *Adv Water Resour* 73:1–15. <https://doi.org/10.1016/j.advwatres.2014.06.010>
- Boergens E, Rangelova E, Sideris MG, Kusche J (2014) Assessment of the capabilities of the temporal and spatio-temporal ICA method for geophysical signal separation in GRACE data. *J Geophys Res Solid Earth* 119:4429–4447. <https://doi.org/10.1002/2013JB010452>
- Broomhead DS, King GP (1986a) Extracting qualitative dynamics from experimental data. *Phys D* 20(2–3):217–236
- Broomhead DS, King GP (1986b) On the qualitative analysis of experimental dynamical systems. *Nonlinear Phenomena Chaos* 113:144
- Cardoso J-F (1999) High-order contrasts for independent component analysis. *Neural Comput* 11:157–192. <https://doi.org/10.1162/089976699300016863>
- Cardoso J-F, Souloumiac A (1993) Blind beamforming for non-Gaussian signals. *IEEE Proc* 140:362–370. <https://doi.org/10.1049/ip-f-2.1993.0054>
- Cardoso J-F, Souloumiac A (1995) Jacobi angles for simultaneous diagonalization. *SIAM J Math Anal Appl* 17:161–164
- Chatfield C (1989) *The analysis of time series: an introduction*. Chapman and Hall/CRC, London
- Comon P (1994a) Independent component analysis: a new concept? *Sig Process* 36(3):287–314
- Comon P (1994b) Tensor diagonalization, a useful tool in signal processing. In: *IFAC symposium on system identification, IFAC-SYSD*, pp 77–82
- Efron B (1979) Bootstrap methods: another look at the Jackknife. *Ann Stat* 7:1–26
- Eicker A, Forootan E, Springer A, Longuevergne L, Kusche J (2016) Does GRACE see the terrestrial water cycle ‘intensifying’? *J Geophys Res Atmos* 121:733–745. <https://doi.org/10.1002/2015JD023808>
- Fasullo JT, Boening C, Landerer FW, Nerem RS (2013) Australia’s unique influence on global sea level in 2010–2011. *Geophys Res Lett* 40:4368–4373. <https://doi.org/10.1002/grl.50834>
- Feldstein SB (2003) The dynamics of NAO teleconnection pattern growth and decay. *QJR Meteorol Soc* 129:901–924. <https://doi.org/10.1256/qj.02.76>
- Fenoglio-Marc L (2001) Analysis and representation of regional sea level variability from altimetry and atmospheric-oceanic data. *Geophys J Int* 145(1):1–18. <https://doi.org/10.1046/j.1365-246x.2001.00284.x>
- Forootan E (2014) *Statistical signal decomposition techniques for analyzing time-variable satellite gravimetry data*. Ph.D. thesis, University of Bonn, Germany. <http://hss.ulb.uni-bonn.de/2014/3766/3766.htm>
- Forootan E, Kusche J (2012) Separation of global time-variable gravity signals into maximally independent components. *J Geod* 86(7):477–497. <https://doi.org/10.1007/s00190-011-0532-5>
- Forootan E, Kusche J (2013) Separation of deterministic signals using independent component analysis (ICA). *Stud Geophys Geod* 57(1):17–26. <https://doi.org/10.1007/s11200-012-0718-1>
- Forootan E, Awange JL, Kusche J, Heck B, Eicker A (2012) Independent patterns of water mass anomalies over Australia from satellite data and models. *Remote Sens Environ* 124:427–443. <https://doi.org/10.1016/j.rse.2012.05.023>
- Forootan E, Kusche J, Loth I, Schuh W-D, Eicker A, Awange J, Longuevergne L, Dieckrueger B, Schmidt M, Shum CK (2014) Multivariate prediction of total water storage anomalies over West Africa from multi-satellite data. *Surv Geophys* 35:913–940. <https://doi.org/10.1007/s10712-014-9292-0S>
- Forootan E, Khandu K, Awange JL, Schumacher M, Anyah RO, van Dijk AIJM, Kusche J (2016) Quantifying the impacts of ENSO and IOD on rain gauge and remotely sensed precipitation products over Australia. *Remote Sens Environ*, vol 172, January 2016, pp 50–66. <https://doi.org/10.1016/j.rse.2015.10.027>

- Frappart F, Ramillien G, Maisongrande P, Bonnet M-P (2010) Denoising satellite gravity signals by independent component analysis. *IEEE Geosci Remote Sens Lett* 7(3):421–425. <https://doi.org/10.1109/LGRS.2009.2037837>
- Frappart F, Ramillien G, Leblanc M, Tweed SO, Bonnet M-P, Maisongrande P (2011a) An independent component analysis filtering approach for estimating continental hydrology in the GRACE gravity data. *Remote Sens Environ* 115(1):187–204. <https://doi.org/10.1016/j.rse.2010.08.017>
- Fu G-S, Anderson M, Adali T (2015) Complex independent component analysis using three types of diversity: non-Gaussianity, nonwhiteness, and noncircularity. *IEEE Trans Signal Process* 63(3):794–805. <https://doi.org/10.1109/TSP.2014.2385047>
- Gualandi A, Serpelloni E, Belardinelli ME (2016) Blind source separation problem in GPS time series. *J Geod* 90:323–341. <https://doi.org/10.1007/s00190-015-0875-4>
- Hannachi A, Jolliffe IT, Stephenson DB (2007) Empirical orthogonal functions and related techniques in atmospheric science: a review. *Int J Climatol* 1152:1119–1152. <https://doi.org/10.1002/joc.1499>
- Hannachi A, Unkel S, Trendafilov NT, Jolliffe IT (2009) Independent component analysis of climate data: a new look at EOF rotation. *J Clim* 22:2797–2812. <https://doi.org/10.1175/2008JCLI2571.1>
- Horel JD (1984) Complex principal component analysis: theory and examples. *J Appl Meteorol* 23(12):1660–1673. [https://doi.org/10.1175/1520-0450\(1984\)023<1660:CPCATA>2.0.CO;2](https://doi.org/10.1175/1520-0450(1984)023<1660:CPCATA>2.0.CO;2)
- Hurrell J W (2003) The North Atlantic Oscillation: climatic significance and environmental impact. American Geophysical Union. ISBN 9780875909943
- Hyvärinen A (1999a) Survey on independent component analysis. *Neural Comput Surv* 2:94–128
- Hyvärinen A (1999b) Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans Neural Netw* 10(3):626–634. <https://doi.org/10.1109/72.761722>
- Hyvärinen A, Oja E (2000) Independent component analysis: algorithms and applications. *Neural Netw* 13(4–5):411–430. [https://doi.org/10.1016/S0893-6080\(00\)00026-5](https://doi.org/10.1016/S0893-6080(00)00026-5)
- James CJ, Hesse CW (2005) Independent component analysis for biomedical signals. *Physiol Meas* 26(1):R15–39. <https://doi.org/10.1088/0967-3334/26/1/R02>
- Jung T-P, Makeig S, McKeown MJ, Bell AJ, Lee T-W, Sejnowski TJ (2005) Imaging brain dynamics using independent component analysis. *Proc IEEE Inst Electr Electron Eng* 89(7):1107–1122
- Koch KR (1999) Parameter estimation and hypothesis testing in linear models, 2nd edn. Springer, New York
- Krishnaswamy J, Vaidyanathan S, Rajagopalan B, Bonell M, Sankaran M, Bhalla RS, Badiger S (2015) Non-stationary and non-linear influence of ENSO and Indian Ocean Dipole on the variability of Indian monsoon rainfall and extreme rain events. *Clim Dyn* 45:175–184. <https://doi.org/10.1007/s00382-014-2288-0>
- Kusche J, Eicker A, Forootan E, Springer A, Longuevergne L (2016) Mapping probabilities of extreme continental water storage changes from space gravimetry. *Geophys Res Lett* 43:8026–8034. <https://doi.org/10.1002/2016GL069538>
- Lian T, Chen D (2012) An evaluation of rotated EOF analysis and its application to tropical Pacific SST variability. *J Clim* 25:5361–5373. <https://doi.org/10.1175/JCLI-D-11-00663.1>
- Liu C, Wechsler H (2003) Independent component analysis of Gabor features for face recognition. *IEEE Trans Neural Netw* 14(4):919–928. <https://doi.org/10.1109/TNN.2003.813829>
- Lorenz EN (1970) Climate change as a mathematical problem. *J Appl Meteorol* 9:325–329
- Matalas NC (1997) Stochastic hydrology in the context of climate change. *Clim Change* 37:89–101
- Ming F, Yang Y, Zeng A et al (2016) Spatiotemporal filtering for regional GPS network in China using independent component analysis. *J Geod*. <https://doi.org/10.1007/s00190-016-0973-y>
- Moore GWK, Halfar J, Majeed H, Adey W, Kronz A (2017) Amplification of the Atlantic Multidecadal Oscillation associated with the onset of the industrial-era warming. *Sci Rep* 7:40861. <https://doi.org/10.1038/srep40861>
- Moreau E (2001) A generalization of joint-diagonalization criteria for source separation. *IEEE Trans Signal Process* 49(3):530–541. <https://doi.org/10.1109/78.905873>
- Omondi P, Awange JL, Ogallo LA, Ininda J, Forootan E (2013) The influence of low frequency sea surface temperature modes on delineated decadal rainfall zones in Eastern Africa region. *Adv Water Resour* 54:161–180. <https://doi.org/10.1016/j.advwatres.2013.01.001>
- Phillips T, Nerem R, Fox-Kemper B, Famiglietti J, Rajagopalan B (2012) The influence of ENSO on global terrestrial water storage using GRACE. *Geophys Res Lett* 39:L16705. <https://doi.org/10.1029/2012GL052495>
- Preisendorfer R (1988) Principal component analysis in meteorology and oceanography. Elsevier, Amsterdam
- Priestley MB (1988) Non-linear and non-stationary time series analysis. Academic Press, London ISBN 0-12-564911-8

- Rangelova E, Sideris M, Kim J (2012) On the capabilities of the multi-channel singular spectrum method for extracting the main periodic and non-periodic variability from weekly GRACE data. *J Geod* 54:64–78. <https://doi.org/10.1016/j.jog.2011.10.006>
- Rasmusson EM, Arkin PA, Chen W-Y, Jalickee JB (1981) Biennial variations in surface temperature over the United States as revealed by singular decomposition. *Mon Weather Rev* 109:587–598. [https://doi.org/10.1175/1520-0493\(1981\)109<0587:BVIStO>2.0.CO;2](https://doi.org/10.1175/1520-0493(1981)109<0587:BVIStO>2.0.CO;2)
- Rencher AC, Christensen WF (2012) Methods of multivariate analysis. Wiley series in probability and statistics, 709, 3rd edn. Wiley, London, p 19 ISBN 9781118391679
- Reynolds RW, Rayner NA, Smith TM, Stokes DC, Wang W (2002) An improved in situ and satellite SST analysis for climate. *J Clim* 15:1609–1625. [https://doi.org/10.1175/1520-0442\(2002\)015<1609:AIISAS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2)
- Richman MB (1986) Rotation of principal components. *J Climatol* 6(3):293–335. <https://doi.org/10.1002/joc.3370060305>
- Saji NH, Goswami BN, Vinayachandran PN, Yamagata T (1999) A dipole mode in the tropical Indian Ocean. *Nature* 401:360–363
- Sawada H, Mukai R, Araki S, Makino S (2005) Frequency-domain blind source separation. In: Speech enhancement, signals and communication technology. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-27489-8_13
- Schmidt R, Petrovic S, Güntner A, Barthelmes F, Wünsch J, Kusche J (2008b) Periodic components of water storage changes from GRACE and global hydrology models. *J Geophys Rese Solid Earth* 113:B08419. <https://doi.org/10.1029/2007JB005363>
- Sharifi MA, Forootan E, Nikkhoo M, Awange J, Najafi M (2013) A point-wise least squares spectral analysis (LSSA) of the Caspian Sea level fluctuations, using Topex/Poseidon and Jason-1 observations. *J Adv Space Res* 51(1):858–873. <https://doi.org/10.1016/j.asr.2012.10.001>
- Shum CK, Kuo C (2010) Observation and geophysical causes of present-day sea level rise. Chapter 7 in Climate Change and Food Security in South Asia. In: Lal R, Sivakumar M, Faiz SMA, Mustafizur Rahman AHM, Islam KR (eds) Springer, Holland
- Talpe MJ, Nerem RS, Forootan E, Schmidt M, Lemoine FG, Enderlin EM, Landerer FW (2017) Ice mass change in Greenland and Antarctica between 1993 and 2013 from satellite gravity measurements. *J Geod.* <https://doi.org/10.1007/s00190-017-1025-y>
- Tapley BD, Bettadpur S, Watkins M, Reigber C (2004) The gravity recovery and climate experiment: mission overview and early results. *Geophys Res Lett* 31:L09607. <https://doi.org/10.1029/2004GL019920>
- Timm O, Pfeiffer M, Dullo W-C (2005) Nonstationary ENSO-precipitation teleconnection over the equatorial Indian Ocean documented in a coral from the Chagos Archipelago. *Geophys Res Lett* 32:L02701. <https://doi.org/10.1029/2004GL021738>
- Trenberth KE (1990) Recent observed interdecadal climate changes in the Northern Hemisphere. *Bull Am Meteorol Soc* 71:988–993. [https://doi.org/10.1175/1520-0477\(1990\)071b0988:ROICCIN2.0.CO;2](https://doi.org/10.1175/1520-0477(1990)071b0988:ROICCIN2.0.CO;2)
- von Storch H, Zwiers F (1999) Statistical analysis in climate research. Cambridge University Press, Cambridge
- Wahr J, Zhong S (2013) Computations of the viscoelastic response of a 3-D compressible Earth to surface loading: an application to Glacial Isostatic Adjustment in Antarctica and Canada. *Geophys J Int* 192:557–572. <https://doi.org/10.1093/gji/ggs030>
- Wallace JM, Smith C, Bretherton CS (1992) Singular value decomposition of wintertime sea surface temperature and 500-mb height anomalies. *J Clim* 5:561–576. [https://doi.org/10.1175/1520-0442\(1992\)005<0561:SVDOWS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1992)005<0561:SVDOWS>2.0.CO;2)
- Weare BC, Nasstrom JN (1982) Examples of extended empirical orthogonal function analyses. *Mon Weather Rev* 110:784–812. [https://doi.org/10.1175/1520-0493\(1982\)110<0481:EEOEOF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1982)110<0481:EEOEOF>2.0.CO;2)
- Westra S, Brown C, Lall U, Sharma A (2007) Modeling multivariable hydrological series: principal component analysis or independent component analysis? *Water Resour Res* 43(6):W06429. <https://doi.org/10.1029/2006WR005617>
- Wouters B, Schrama EJO (2007) Improved accuracy of GRACE gravity solutions through empirical orthogonal function filtering of spherical harmonics. *Geophys Res Lett* 34:L23711. <https://doi.org/10.1029/2007GL032098>
- Wu Y, Wu B, Liu J, Lu H (2008) Probabilistic tracking on Riemannian manifolds. In: IEEE 19th international conference on pattern recognition. <https://doi.org/10.1109/ICPR.2008.4761046>